GRATINGS: THEORY AND NUMERIC APPLICATIONS

Tryfon Antonakakis
Fadi Baïda
Abderrahmane Belkhir
Kirill Cherednichenko
Shane Cooper
Richard Craster
Guillaume Demesy
John DeSanto
Gérard Granet
Boris Gralak
Sébastien Guenneau
Daniel Maystre
André Nicolet
Brian Stout
Gérard Tayeb
Fréderic Zolla
Benjamin Vial

Evgeny Popov, Editor

Institut Fresnel, Université d'Aix-Marseille, Marseille, France
Femto, Université de Franche-Compté, Besançon, France
LASMEA, Université Blaise Pascal, Clermont-Ferrand, France
Colorado School of Mines, Golden, USA
CERN, Geneva, Switzerland
Imperial College London, UK
Cardiff University, Cardiff, UK
Mouloud Mammeri Univesity, Tizi-Ouzou, Algeria

ISBN: 2-85399-860-4

www.fresnel.fr/numerical-grating-book

ISBN: 2-85399-860-4

First Edition, 2012

World Wide Web:

www.fresnel.fr/numerical-grating-book

Institut Fresnel, Université d'Aix-Marseille, CNRS
Faculté Saint Jérôme,
13397 Marseille Cedex 20,
France

Gratings: Theory and Numeric Applications, Evgeny Popov, editor (Institut Fresnel, CNRS, AMU, 2012)

Chapter 7:

Differential Theory of Periodic Structures

Evgeny Popov

Table of Contents:

Chapter 7

Differential Theory of Periodic Structures

Evgeny Popov

Institut Fresnel, CNRS, Aix-Marseille Université,
Campus de Saint Jerome,13013 Marseille, France
e.popov@fresnel.fr *www.fresnel.fr/perso/popov*

The basic idea of the differential methods consisits in projecting the electromagnetic field on a set of basic functions in order to reduce Maxwell partial differential equations into a set of ordinary differential equations. When working in a Cartesian coordinates, the natural basis consists of exponentials, using the periodicity of the optogeometrical parameters. Diffraction by a single aperture requires working in the basis of cylindrical Bessel functions [7.1], while diffraction by an arbitrary-shaped single object requires vector spherical functions [7.2] as a basis.

The first studies using the differential method [7.3] appeared in the late 1960s, initiated by the birth of the computers. These studies concerned the modeling of diffusion of particles in nuclear potential by using the separation of variables of the radial Schrödinger equation. The method was called "optical method" due to the similarity between the Schrödinger and the Helmholtz equations. The first applications to grating diffraction appear in 1969 [7.4], but accurate and converging results required combining the differential method with conformal mapping techniques [7.5]. The classical differential theory as known nowadays was formulated in [7.6, 7.7]. One can find a detailed review on the classical differential method in [7.8]

It appeared that the classical differential theory suffered from severe numerical problems in transverse magnetic (TM) polarization, as well as for deep gratings. The first breakthrough was made in the first half of the 1990s, by introducing orthonormalization of the differential equations during their integration [7.9] and followed later by the so-called R-matrix or S-matrix propagating algorithms [7.10]. The second breakthrough improved considerably the convergence in TM polarization for lamellar gratings, by introducing the correct factorization rules (see further on), at first by chance [7.11] and after that using theoretical arguments [7.12], closely followed by a generalization to arbitrary profiles [7.13]. A detailed review can be found in [7.14].

7.1. Maxwell equations in the truncated Fourier space

Let us consider a structure with two-dimensional periodicity along the x- and y-axis (Fig.7.1) with periods equal to d_x and d_y. The modulated (grating) region extends in z from z_{min} to z_{max}. Inside that region, for a given value of the vertical coordinate z, the permittivity ε and permeability μ are periodic functions in x and y that can be projected on exponential Fourier basis:

$$\varepsilon(x,y,z)=\sum_{m,n=-\infty}^{+\infty}\varepsilon_{m,n}(z)\exp(imK_x x+inK_y y)$$
$$\mu(x,y,z)=\sum_{m,n=-\infty}^{+\infty}\mu_{m,n}(z)\exp(imK_x x+inK_y y) \qquad (7.1)$$

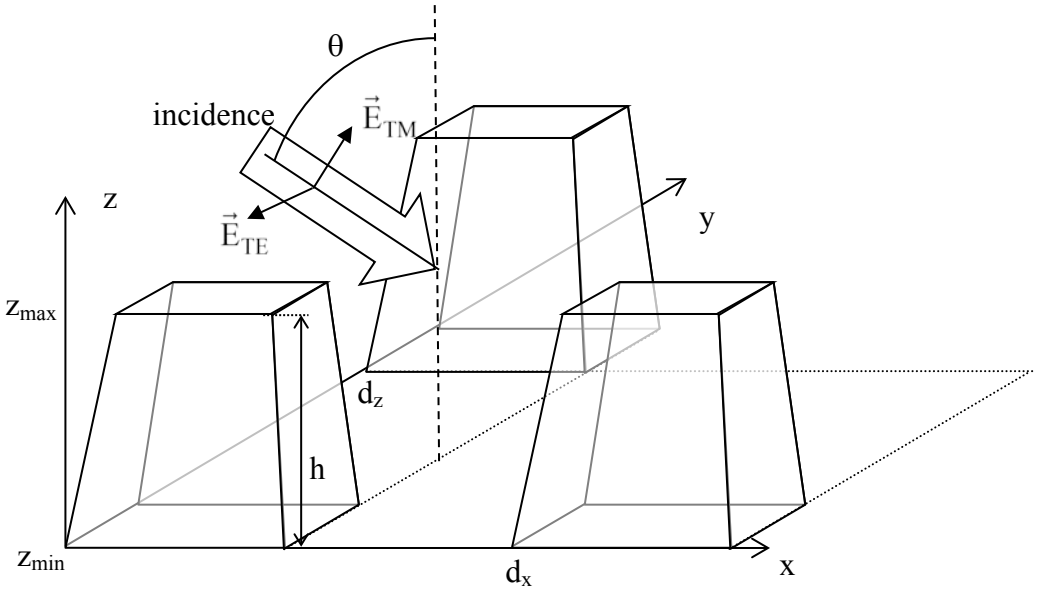


Fig.7.1. Schematical representation of a structure having two-dimensional periodicity in x and y-directions, consisting of truncated pyramids with height h.

where $K_x=2\pi/d_x$ and $K_y=2\pi/d_y$. We shall deal with a monochromatic (wavelength λ) plane wave incident on the structure with a wavevector:

$$\vec{k}_{inc}=(\alpha_0,\beta_0,-\gamma_0) \qquad (7.2)$$

with components related to the incident polar angle θ (between the incident direction and the grating normal) and azimuthal angle φ (between the plane of incidence and the xOz-plane):

$$\alpha_0=k_0\sin\theta\cos\varphi,\quad \beta_0=k_0\sin\theta\cos\varphi,$$
$$\gamma_0=\sqrt{k_0^2n_{inc}^2-\alpha_0^2-\beta_0^2},\quad k_0=2\pi/\lambda \qquad (7.3)$$

where n_{inc} is the refractive index of the cladding.

The existence and uniqueness of the solution of the diffraction problem is an interesting problem that is not discussed here. The reader can refer to several basic works (see for example [7.15, 7.16]. What is important to conclude is that the electromagnetic field is pseudo-periodic, so that similarly to eq.(7.1), the electric $\vec{E}$ and magnetic $\vec{H}$ field vectors can be represented in pseudo-Fourier series:

$$\vec{E}(x,y,z)=\sum_{m,n=-\infty}^{+\infty}\vec{E}_{m,n}(z)\exp\left[i(\alpha_0+mK_x)x+i(\beta_0+nK_y)y\right]$$
$$\vec{H}(x,y,z)=\sum_{m,n=-\infty}^{+\infty}\vec{H}_{m,n}(z)\exp\left[i(\alpha_0+mK_x)x+i(\beta_0+nK_y)y\right] \tag{7.4}$$

In what follows, we use the notations:

$$\alpha_m=\alpha_0+mK_x,\quad \beta_n=\beta_0+nK_y. \tag{7.5}$$

From a numerical point of view, it is necessary to truncate the series in eqs.(7.1) and (7.4), introducing truncation parameters N_x and N_y, which limit the lower and the upper boundaries in the series.

Maxwell equations written in Fourier space take the form, assuming $\exp(-i\omega t)$ time dependence with circular frequency ω:

$$\begin{aligned}
&i\beta_{m,n}E_{z,m,n}(z)-\frac{d}{dz}E_{y,m,n}(z)=i\omega B_{x,m,n}(z)\\
&\frac{d}{dz}E_{x,m,n}(z)-i\alpha_{m,n}E_{z,m,n}(z)=i\omega B_{y,m,n}(z)\\
&i\alpha_{m,n}E_{y,m,n}(z)-i\beta_{m,n}E_{x,m,n}(z)=i\omega B_{z,m,n}(z)\\
&i\beta_{m,n}H_{z,m,n}(z)-\frac{d}{dz}H_{y,m,n}(z)=-i\omega D_{x,m,n}(z)\\
&\frac{d}{dz}H_{x,m,n}(z)-i\alpha_{m,n}H_{z,m,n}(z)=-i\omega D_{y,m,n}(z)\\
&i\alpha_{m,n}H_{y,m,n}(z)-i\beta_{m,n}H_{x,m,n}(z)=-i\omega D_{z,m,n}(z).
\end{aligned} \tag{7.6}$$

As can be observed, the third and the sixth equations are not differential equations, and they are used to eliminate the z-components of the fields, as shown further on. It has to be stressed out that the equations with different (m,n) numbers are coupled through the Fourier components of $\vec{D}=\varepsilon\vec{E}$ and $\vec{B}=\mu\vec{H}$.

The next step is to factorize the products $\vec{D}=\varepsilon\vec{E}$ and $\vec{B}=\mu\vec{H}$. In this chapter we assume media with linear dielectric and magnetic properties and without spontaneous polarizations. The problem of Fourier transform of the product of two functions

$$\vec{D}(x,y,z)=\sum_{m,n=-\infty}^{+\infty}\vec{D}_{m,n}(z)\exp\left[i(\alpha_0+mK_x)x+i(\beta_0+nK_y)y\right] \tag{7.7}$$

is, in generally, solved theoretically by convolution of the Fourier transformers of the two functions, using the so-called Laurent's rule:

$$\vec{D}_{m,n}(z)=\sum_{m',n'=-\infty}^{+\infty}\varepsilon_{m-m',n-n'}(z)\vec{E}_{m',n'}(z). \tag{7.8}$$

However, there are several problems in the numerical application of this rule:

First, numerical applications are simplified when using matrix notations. However, most of the standard routines use single-rank vectors and rectangular (2-ranks) matrices, while the vectors D and E in eq.(7.8) have two indexes, and the matrix ε depend on four indexes. In the case of classical grating with one-dimensional periodicity, this problem does not exist. Fortunately, for structures having 2D periodicity, a reduction to standard arrays is possible by introduction of a single index instead of the double for the vectors, by the following substitution:

$$p = (m + N_x)(2N_y + 1) + (n + N_y + 1) \tag{7.9}$$

so that when m varies between $-N_x$ and $+N_x$ and n varies between $-N_y$ and $+N_y$, p varies between 1 and $P_{max} = (2N_x+1)(2N_y+1)$. Using these notations, we can introduce standard arrays in the following manner:

$$\begin{aligned} &\vec{D}_p(z) = \vec{D}_{m,n}(z), \quad \vec{E}_p(z) = \vec{E}_{m,n}(z), \quad \text{etc. for } \vec{H} \text{ and } \vec{B}, \\ &\varepsilon_{p-p'}(z) = \varepsilon_{m-m',n-n'}(z) \end{aligned} \tag{7.10}$$

so that eq.(7.8) takes the standard truncated form

$$\vec{D}_p(z) = \sum_{p'=1}^{P_{max}} \varepsilon_{p-p'}(z)\vec{E}_{p'}(z). \tag{7.11}$$

That can be written in matrix notations in the form

$$\left[\vec{D}(z)\right] = \left[\!\left[\varepsilon(z)\right]\!\right]\left[\vec{E}(z)\right], \tag{7.12}$$

where double square brackets stand for the Toepliz matrix.

In addition, two diagonal matrices are useful:

$$\begin{aligned} \alpha_{p,p'} &= \delta_{p,p'}\alpha_m \\ \beta_{p,p'} &= \delta_{p,p'}\beta_n \end{aligned}, \tag{7.13}$$

with $\delta_{p,p'}$ being the Kronecker's symbol.

Second, due to the vectorial character of the fields, the matrix form in eq.(7.12) has to be interpreted in a block form:

$$\begin{pmatrix} [D_x(z)] \\ [D_y(z)] \\ [D_z(z)] \end{pmatrix} = \left[\!\left[\varepsilon(z)\right]\!\right] \begin{pmatrix} [E_x(z)] \\ [E_y(z)] \\ [E_z(z)] \end{pmatrix}, \text{ isotropic media} \tag{7.14}$$

$$\begin{pmatrix} [D_x(z)] \\ [D_y(z)] \\ [D_z(z)] \end{pmatrix} = \begin{pmatrix} \left[\!\left[\varepsilon_{xx}(z)\right]\!\right] & \left[\!\left[\varepsilon_{xy}(z)\right]\!\right] & \left[\!\left[\varepsilon_{xz}(z)\right]\!\right] \\ \left[\!\left[\varepsilon_{yx}(z)\right]\!\right] & \left[\!\left[\varepsilon_{yy}(z)\right]\!\right] & \left[\!\left[\varepsilon_{yz}(z)\right]\!\right] \\ \left[\!\left[\varepsilon_{zx}(z)\right]\!\right] & \left[\!\left[\varepsilon_{zy}(z)\right]\!\right] & \left[\!\left[\varepsilon_{zz}(z)\right]\!\right] \end{pmatrix} \begin{pmatrix} [E_x(z)] \\ [E_y(z)] \\ [E_z(z)] \end{pmatrix}, \text{ anisotropic media.} \tag{7.15}$$

The **third** problem linked with the truncation of eq. (7.8) has limited the use of the differential methods (including RCW method) for more than 30 years, and has been solved for lamellar gratings in the late 90s [7.11, 7.12], and for arbitrary-profile gratings in the start of the 2000s [7.13]. The problem is due to the very slow convergence with respect to the number of Fourier components in the truncated sum of eq. (7.8), when the two functions in the product are discontinuous. As demonstrated by Li [7.12], four different cases can be distinguished with respect to eq.(7.12):

1. Both ε and E are continuous functions of x and y.
2. ε is discontinuous, but E is continuous. This is the case of the tangential component of E.
3. Both ε and E are discontinuous, but their product D is continuous, as it happens for the normal component of D.
4. All three functions are discontinuous.

In the first and second case, Laurent's rule assures relatively rapid convergence. In the third case, more rapidly converging scheme can be obtained through the following considerations for isotropic media.

If D = εE is continuous, then it is possible to factorize the product between D (continuous) and 1/ε (discontinuous) using the Laurent's rule (called by Li *direct* rule):

$$\left[\vec{E}(z)\right] = \left[\!\left[\frac{1}{\varepsilon(z)}\right]\!\right]\left[\vec{D}(z)\right], \tag{7.16}$$

wherefrom the so called *inverse* rule is formulated:

$$\left[\vec{D}(z)\right] = \left[\!\left[\frac{1}{\varepsilon(z)}\right]\!\right]^{-1}\left[\vec{E}(z)\right], \tag{7.17}$$

which can be applied if the matrix $\left[\!\left[\frac{1}{\varepsilon(z)}\right]\!\right]$ is not singular, a requirement that can create numerical problem for highly conducting gratings having small imaginary part of ε.

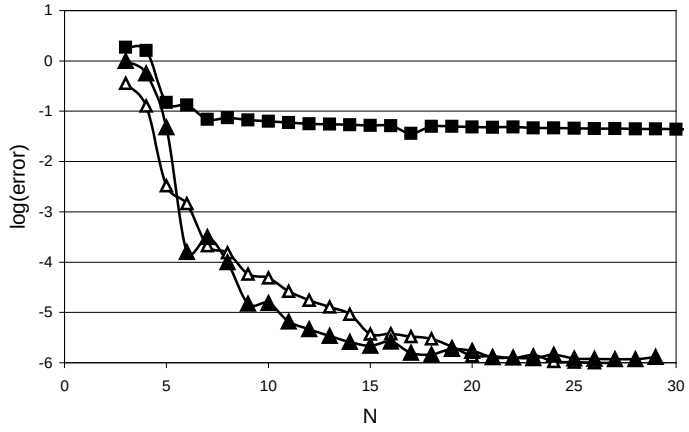


Fig.7.2.Convergence of the classical and the FFF version of the differential theory in the case of a dielectric sinusoidal grating with high contrast. Squares, old version of the differential theory for TM polarization; open triangles, new version, TM polarization; solid triangles, TE polarization (after [7.13]).

When infinite series are considered, eq.(7.17) is identical with eq.(7.12). However, as shown in Fig.7.2, the correct use of the direct or the inverse rules improves drastically the convergence of the differential methods with respect to the truncation parameter. Similarly to the abbreviation FFT, standing for Fast Fourier transformation, we have introduced the term Fast Fourier factorization (FFF) to name the correct use of the direct and the inverse rules, when applied numerically in the truncated Fourier space.

In the fourth case, neither the direct, nor the inverse rule result in acceptable convergence, so that this case must be avoided. Fortunately, this can be done by considering separately the electromagnetic field components, tangential or normal to the grating profile and taking into account that the electric field components tangential to the surface separating two different permittivities are continuous, in the same way as the normal components of the displacement D.

7.2. Differential theory for crossed gratings made of isotropic materials

In the isotropic case, the displacement vector $\vec{D}$ can easily be separated into a continuous part $\vec{D}_N = \varepsilon\vec{E}_N$, normal to the profile surface, and $\vec{D}_T = \varepsilon\vec{E}_T$ that contains the continuous function $\vec{E}_T$. Let us define a unit vector $\vec{N}$, normal to the grating profile. Although it is well defined on the profile (except edges), it is necessary to generalize its definition all over the grating region, which cannot be done in a unique manner. Different choices are shown further on for specific gratings having 1D or 2D periodicity. Using this generalized vector, the relations between $\vec{E}$ and $\vec{D}$ can be decomposed into two terms, for each of which we can apply the direct or the inverse factorization rules, skipping the explicit writing of the z-dependence:

$$\vec{D} = \varepsilon\vec{E}_N + \varepsilon\vec{E}_T = \varepsilon\vec{N}\left(\vec{N}\cdot\vec{E}\right) + \varepsilon\left[\vec{E} - \vec{N}\left(\vec{N}\cdot\vec{E}\right)\right]. \tag{7.18}$$

The first term is a product of type 2 and requires the direct rule. The second term is of type 3, demanding the inverse rule, so that:

$$\begin{aligned}\left[\vec{D}\right] &= \left[\!\left[\varepsilon\right]\!\right]\left[\vec{E}_T\right] + \left[\!\left[\frac{1}{\varepsilon}\right]\!\right]^{-1}\left[\vec{E}_N\right] \\ &= \left[\!\left[\varepsilon\right]\!\right]\left[\vec{E} - \vec{N}\left(\vec{N}.\vec{E}\right)\right] + \left[\!\left[\frac{1}{\varepsilon}\right]\!\right]^{-1}\left[\vec{N}\left(\vec{N}.\vec{E}\right)\right].\end{aligned} \tag{7.19}$$

Introducing a square matrix representing a tensor product denoted $\left(\vec{N}\vec{N}\right)$ with elements given by N_iN_j, we obtain:

$$\left[\vec{D}\right] = \left[\!\left[\varepsilon\right]\!\right]\left[\vec{E}\right] + \left(\left[\!\left[\frac{1}{\varepsilon}\right]\!\right]^{-1} - \left[\!\left[\varepsilon\right]\!\right]\right)\left[\!\left[\vec{N}\vec{N}\right]\!\right]\left[\vec{E}\right] = Q_\varepsilon\left[\vec{E}\right], \tag{7.20}$$

where the matrix Q_ε has the form:

$$Q_\varepsilon = \left[\!\left[\varepsilon\right]\!\right] + \left(\left[\!\left[\frac{1}{\varepsilon}\right]\!\right]^{-1} - \left[\!\left[\varepsilon\right]\!\right]\right)\left[\!\left[\vec{N}\vec{N}\right]\!\right]. \tag{7.21}$$

In a similar manner for magnetic materials, we can find the link between magnetic field and induction in the truncated Fourier space:

$$\left[\vec{B}\right] = Q_\mu \left[\vec{H}\right],$$
$$\text{with} \quad Q_\mu = Q_\varepsilon = [\![\mu]\!] + \left([\![\frac{1}{\mu}]\!]^{-1} - [\![\mu]\!] \right) [\![\vec{N}\vec{N}]\!] \tag{7.22}$$

Eq.(7.20) allows eliminating E_z in the system (7.6):

$$\begin{aligned} \left[E_z\right] &= Q_{\varepsilon,zz}^{-1} \left(\left[D_z\right] - Q_{\varepsilon,zx} \left[E_x\right] - Q_{\varepsilon,zy} \left[E_y\right] \right) \\ &= -Q_{\varepsilon,zz}^{-1} \left(\frac{\alpha \left[H_y\right] - \beta \left[H_x\right]}{\omega} + Q_{\varepsilon,zx} \left[E_x\right] + Q_{\varepsilon,zy} \left[E_y\right] \right) \end{aligned} \tag{7.23}$$

where the matrices α and β are defined in eq.(7.13).

Repeating the procedure for H_z:

$$\left[H_z\right] = Q_{\mu,zz}^{-1} \left(\frac{\alpha \left[E_y\right] - \beta \left[E_x\right]}{\omega} - Q_{\mu,zx} \left[H_x\right] - Q_{\mu,zy} \left[H_y\right] \right), \tag{7.24}$$

it is also eliminated from eqs. (7.6).

For **non-magnetic** media, the last expression is further simplified:

$$\left[H_z\right] = \frac{\alpha \left[E_y\right] - \beta \left[E_x\right]}{\omega \mu_0}. \tag{7.25}$$

Thus the system (7.6) is replaced by a system of ordinary differential equations:

$$\frac{d}{dz} \begin{pmatrix} \left[E_x\right] \\ \left[E_y\right] \\ \left[H_x\right] \\ \left[H_y\right] \end{pmatrix} = iM \begin{pmatrix} \left[E_x\right] \\ \left[E_y\right] \\ \left[H_x\right] \\ \left[H_y\right] \end{pmatrix}. \tag{7.26}$$

This equation can be expressed in a compressed form:

$$\frac{d}{dz} F(z) = iM(z) F(z) \tag{7.27}$$

Here the matrix M has 4x4 blocks:

$$
\begin{aligned}
M_{11} &= -\alpha Q_{\varepsilon,zz}^{-1} Q_{\varepsilon,zx} - Q_{\mu,yz} Q_{\mu,zz}^{-1} \beta \\
M_{12} &= -\alpha Q_{\varepsilon,zz}^{-1} Q_{\varepsilon,zy} + Q_{\mu,yz} Q_{\mu,zz}^{-1} \alpha \\
M_{13} &= -\omega Q_{\mu,yz} Q_{\mu,zz}^{-1} Q_{\mu,zx} + \frac{\alpha}{\omega} Q_{\varepsilon,zz}^{-1} \beta + \omega Q_{\mu,yx} \\
M_{14} &= -\omega Q_{\mu,yz} Q_{\mu,zz}^{-1} Q_{\mu,zy} - \frac{\alpha}{\omega} Q_{\varepsilon,zz}^{-1} \alpha + \omega Q_{\mu,yy} \\
M_{21} &= -\beta Q_{\varepsilon,zz}^{-1} Q_{\varepsilon,zx} + Q_{\mu,xz} Q_{\mu,zz}^{-1} \beta \\
M_{22} &= -\beta Q_{\varepsilon,zz}^{-1} Q_{\varepsilon,zy} - Q_{\mu,xz} Q_{\mu,zz}^{-1} \alpha \\
M_{23} &= \omega Q_{\mu,xz} Q_{\mu,zz}^{-1} Q_{\mu,zx} + \frac{\beta}{\omega} Q_{\varepsilon,zz}^{-1} \beta - \omega Q_{\mu,xx} \\
M_{24} &= \omega Q_{\mu,xz} Q_{\mu,zz}^{-1} Q_{\mu,zy} - \frac{\beta}{\omega} Q_{\varepsilon,zz}^{-1} \alpha - \omega Q_{\mu,xy} \\
M_{31} &= \omega Q_{\varepsilon,yz} Q_{\varepsilon,zz}^{-1} Q_{\varepsilon,zx} - \frac{\alpha}{\omega} Q_{\mu,zz}^{-1} \beta - \omega Q_{\varepsilon,yx} \\
M_{32} &= \omega Q_{\varepsilon,yz} Q_{\varepsilon,zz}^{-1} Q_{\varepsilon,zy} + \frac{\alpha}{\omega} Q_{\mu,zz}^{-1} \alpha - \omega Q_{\varepsilon,yy} \\
M_{33} &= -\alpha Q_{\mu,zz}^{-1} Q_{\mu,zx} - Q_{\varepsilon,yz} Q_{\varepsilon,zz}^{-1} \beta \\
M_{34} &= -\alpha Q_{\mu,zz}^{-1} Q_{\mu,zy} + Q_{\varepsilon,yz} Q_{\varepsilon,zz}^{-1} \alpha \\
M_{41} &= -\omega Q_{\varepsilon,xz} Q_{\varepsilon,zz}^{-1} Q_{\varepsilon,zx} - \frac{\beta}{\omega} Q_{\mu,zz}^{-1} \beta + \omega Q_{\varepsilon,xx} \\
M_{42} &= -\omega Q_{\varepsilon,xz} Q_{\varepsilon,zz}^{-1} Q_{\varepsilon,zy} + \frac{\beta}{\omega} Q_{\mu,zz}^{-1} \alpha + \omega Q_{\varepsilon,xy} \\
M_{43} &= -\beta Q_{\mu,zz}^{-1} Q_{\mu,zx} + Q_{\varepsilon,xz} Q_{\varepsilon,zz}^{-1} \beta \\
M_{44} &= -\beta Q_{\mu,zz}^{-1} Q_{\mu,zy} - Q_{\varepsilon,xz} Q_{\varepsilon,zz}^{-1} \alpha .
\end{aligned}
\tag{7.28}
$$

This form looks like the form of the M-matrix obtained by Lifeng Li for crossed anisotropic (electrically and magnetically) gratings with profiles invariant with respect to z [7.17].

Whatever the form of the matrix M, eq.(7.26) represents a linear set of first-order ordinary differential equations. It can be solved numerically (with several problems, discussed further on), using well developed numerical schemes. In the case of vertical invariance of the optogeometrical parameters of the system inside the modulated region, the elements of the M-matrix becomes constant in z, so that the solution of eq. (7.26) can be found through the eigenvectors and eigenvalues of M, a technique known under the name of Fourier modal method, or Rigorous coupled wave (RCW) method.

The solution of (7.26) gives a linear link between the field in the substrate and in the cladding

$$F(z_{max}) = T\, F(z_{min}), \tag{7.29}$$

where T is called transmission matrix.

The advantage of this presentation comes from the fact that the field components participating in the calculations are tangential to the interfaces between the substrate and the modulated region, and between the cladding and the modulated region, so that they are continuous across these interfaces (in the absence of surface charges).

7.3. Electromagnetic field in the homogeneous regions – plane wave expansion

In most case, the substrate and cladding are homogeneous isotropic media. The electromagnetic field there can be expressed as a sum of plane waves. In particular, if the x and y-dependencies are given as in eq.(7.7), the z-dependence is explicitly known, for example for the electric field it takes the form:

$$\vec{E}_p(z) = \vec{A}_p^{+} \exp(i\gamma_p z) + \vec{A}_p^{-} \exp(-i\gamma_p z) \tag{7.30}$$

of two waves propagating upwards (sign +) and downwards (sign –) along the z-axis, with p given in eq.(7.9). Each diffraction order with a given p propagates independently of the others, the coupling is effective inside the grating region.

The z-propagation constant γ depends on the medium properties:

$$\gamma_p = \sqrt{\omega^2 \varepsilon\mu - \alpha_p^2 - \beta_p^2}\,. \tag{7.31}$$

Equations (7.6) enable us to express the magnetic field components through the electric ones:

$$\begin{aligned} H_{x,p} &= -\frac{1}{\pm\gamma_p}\left(\frac{\alpha_p\beta_p}{\omega\mu}E_{x,p} + \frac{\beta_p^2+\gamma_p^2}{\omega\mu}E_{y,p}\right) \\ H_{y,p} &= \frac{1}{\pm\gamma_p}\left(\frac{\alpha_p^2+\gamma_p^2}{\omega\mu}E_{x,p} + \frac{\alpha_p\beta_p}{\omega\mu}E_{y,p}\right) \end{aligned} \tag{7.32}$$

where the sign of γ determines the direction of propagation in along z-axis.

With this link in mind, the column vector F in eq.(7.27) takes the form:

$$F \equiv \begin{pmatrix} [E_x] \\ [E_y] \\ [H_x] \\ [H_y] \end{pmatrix} = \Psi^{+}A^{+} + \Psi^{-}A^{-}, \tag{7.33}$$

where the column vectors

$$A^{\pm} = \begin{pmatrix} [A_x^{\pm}] \\ [A_y^{\pm}] \end{pmatrix} \tag{7.34}$$

contains the amplitudes of E_x and E_y propagating in positive or negative direction of the z-axis, matrices $\Psi^{\pm}$ are block-diagonal:

$$\Psi^{\pm} = \begin{pmatrix} \mathbb{I} & \\ & \mathbb{I} \\ \Psi^{\pm}_{xx} & \Psi^{\pm}_{xy} \\ \Psi^{\pm}_{yx} & \Psi^{\pm}_{yy} \end{pmatrix}, \tag{7.35}$$

with diagonal blocks

$$\begin{aligned} &\mathbb{I}_{pp} = 1, \\ &\Psi^{\pm}_{xx,pp} = \mp\frac{\alpha_p\beta_p}{\gamma_p\omega\mu}, \quad \Psi^{\pm}_{xy,pp} = \mp\frac{\beta_p^2+\gamma_p^2}{\gamma_p\omega\mu} \\ &\Psi^{\pm}_{yx,pp} = \pm\frac{\alpha_p^2+\gamma_p^2}{\gamma_p\omega\mu}, \quad \Psi^{\pm}_{yy,pp} = \pm\frac{\alpha_p\beta_p}{\gamma_p\omega\mu} \end{aligned} \tag{7.36}$$

found from eq.(7.32)

Let us consider the case of a single incident wave from the cladding. The grating generates different diffraction order that propagate upwards in the cladding and downwards in the substrate. We attribute number 1 to the substrate and number 3 to the cladding. The total number of unknown diffracted field amplitudes will be equal to $4P_{max}$, two sets of $A^{1-}_{x,p}$ and $A^{1-}_{y,p}$ transmitted in the substrate, and two sets of $A^{3+}_{x,p}$ and $A^{3+}_{y,p}$. These unknown amplitudes are subjected to $4P_{max}$ number of linear algebraic equation in (7.29).

In order to obtain the T-matrix, the numerical integration of eq.(7.26) is made by using the so-called shooting method, which consists of choosing $2P_{max}$ linearly independent representatives of the transmitted field. These representatives must correctly reflect the link between the electric and magnetic field components, as given by eqs.(7.32). A typical example for the shooting vectors starting from the substrate is that matrix Ψ^{1+}, which has $2P_{max}$ linearly independent columns. Here again, the number 1 indicates the substrate.

Thus the F column vector at $z = z_{min}$ can be formally written as a linear combination of the unknown amplitudes A^{1-}:

$$\tilde{F}(z_{min}) = \Psi^{1-}A^{1-}, \tag{7.37}$$

Assuming that there is no incidence from the substrate side. Here the tilde indicates that the vector F is not yet the true solution of the diffraction problem.

The result of the numerical integration from z_{min} to z_{max} will provide the values of $\tilde{F}$ at $z = z_{max}$, which are also a linear combination $\tilde{F}(z_{max})A^{1-}$ of A^{1-}, due to the linearity of the problem. On the other side, the column vector F at the upper interface is equal to $\psi^{3+}A^{3+} + \psi^{3-}A^{3-}$, according eq.(7.33), thus a linear set of algebraic equations for the unknown amplitudes A^{1-} and A^{3+} is obtained, with the free part determined by the wave incident from the cladding side:

$$\psi^{3+}A^{3+} + \psi^{3-}A^{3-} = \tilde{F}(z_{max})A^{1+}. \tag{7.38}$$

Once this system is solved, all field components can be calculated.

Unfortunately, this simple procedure creates enormous numerical problems that can be explained by using two different arguments:

First, it is known (but no quite well) in the theory of systems of ordinary differential equations, that numerical integration could become instable after a specific integration length, due to the fact that the set of shooting vectors can lose its linear independence during the integration. In other words, if the initial choice covers a vector space of $2P_{max}$ dimensions, this space could shrink during the numerical integration to reduce its dimensions, so that the final algebraic system (7.38) could become singular. A solution of the problem based on this understanding was proposed in 1990 by G. Tayeb by using intermediate orthonormalization procedures during the numerical integration.

The second argument is based on the fact that inside the modulated region, as well as in the homogeneous regions, electromagnetic field contains components that propagate both in the positive and in the negative z-direction. During the integration, they both are treated in the same manner. As far as the solution requires taking into account the evanescent orders in addition to the propagating ones, a part of the former grows exponentially in z-direction, while the other part decreases exponentially. Due to the limited length of computer words, the ones that decrease substantially will be lost with respect to the ones that grow rapidly, even if the former could bring physical information. During the 90s, several different algorithms were proposed for solving the problem, based on a different treatment of the diffraction orders propagating upwards and downwards [7.9, 7.10]. Among them, to so called S-matrix propagation algorithm [7.10c] is probably the easiest to implement. Moreover, it can be used with methods other than the differential one in, for example, treating a stack of layers by the integral method, or by methods based on a transformation of the coordinate system. Interested reader can find in Appendix 7.1 a brief description of the S-matrix algorithm.

7.4. Several simpler isotropic cases

In practice most applications use non-magnetic materials, for which the form of M-matrix is considerably simplified, taking into account that then Q_μ is diagonal and equal to μ_0. Furthermore, several specific cases are of great interest for application, and they lead to a further simplification of the M-matrix.

7.4.1. Classical grating with one-dimensional periodicity, example of a sinusoidal profile

Let us consider a classical grating with grooves parallel to the y-axis and surface profile given by the equation z = g(x). The vector normal to the surface is given by

$$\begin{aligned} \vec{N} &= \frac{1}{\sqrt{1+g'^2(x)}}\left(-g'(x), 0, 1\right), \quad \text{if } g'(x) \text{ exists,} \\ \vec{N} &= (1,0,0), \quad \text{if not} \end{aligned} \qquad . \tag{7.39}$$

where the prime stands for a derivative with respect to x. In case of vertical walls $\vec{N} = (1, 0, 0)$. Thus the easiest way to generalize the normal vector to the entire modulated region is just to make it equal to eq.(7.39) not only on the profile z = g(x), but everywhere inside the grating region for $\min[g(x)] \le z \le \max[g(x)]$. The advantage of this choice is that $\vec{N}$ does not depend on z, and the Fourier transformation of the tensor $\vec{N}\vec{N}$ is done only once.

If the derivative of the profile function does not exist, or if the function is a multivalued one (e.g., circular or elliptical rods), but the interface can be expressed as a two-variable function:

$$g(x,y)=0\,, \tag{7.40}$$

the normal vector is easily defined as the gradient of the profile function $\vec{N}=\mathrm{grad}\left[g(x,z)\right]\Big/\left|\mathrm{grad}\left[g(x,z)\right]\right|$.

The Q_ε matrix takes the form:

$$Q_\varepsilon=\begin{pmatrix} [\![\varepsilon]\!]\left[\!\left[N_z^2\right]\!\right]+\left[\!\left[\frac{1}{\varepsilon}\right]\!\right]^{-1}\left[\!\left[N_x^2\right]\!\right] & 0 & \left(\left[\!\left[\frac{1}{\varepsilon}\right]\!\right]^{-1}-[\![\varepsilon]\!]\right)[\![N_xN_z]\!] \\ 0 & [\![\varepsilon]\!] & 0 \\ \left(\left[\!\left[\frac{1}{\varepsilon}\right]\!\right]^{-1}-[\![\varepsilon]\!]\right)[\![N_xN_z]\!] & 0 & [\![\varepsilon]\!]\left[\!\left[N_x^2\right]\!\right]+\left[\!\left[\frac{1}{\varepsilon}\right]\!\right]^{-1}\left[\!\left[N_z^2\right]\!\right] \end{pmatrix} \tag{7.41}$$

where it is taken into account that $N_x^2+N_z^2=1$. The fact that the normal vector components participate in the form of products in couples is important, because it leads to the conclusion is that the choice of the sign of $\vec{N}$ plays no role.

Further simplification of the M-matrix comes if limited to non-conical diffraction with $\beta_0 = 0$:

$$M=\begin{pmatrix} -\alpha Q_{\varepsilon,zz}^{-1}Q_{\varepsilon,zx} & 0 & 0 & -\frac{\alpha}{\omega}Q_{\varepsilon,zz}^{-1}\alpha+\omega\mu_0\mathbb{I} \\ 0 & 0 & -\omega\mu_0\mathbb{I} & 0 \\ 0 & \frac{\alpha\alpha}{\omega\mu_0}-\omega[\![\varepsilon]\!] & 0 & 0 \\ -\omega Q_{\varepsilon,xz}Q_{\varepsilon,zz}^{-1}Q_{\varepsilon,zx}+\omega Q_{\varepsilon,xx} & 0 & 0 & -Q_{\varepsilon,xz}Q_{\varepsilon,zz}^{-1}\alpha \end{pmatrix} \tag{7.42}$$

This shows that the system to integrate decouples into two subsystems, corresponding to the two fundamental polarizations, transversal with respect to the plane of incidence, transverse electric (TE):

$$\begin{aligned} \frac{d}{dz}\left[E_y\right]&=-i\omega\mu_0\left[H_x\right] \\ \frac{d}{dz}\left[H_x\right]&=i\left(\frac{\alpha^2}{\omega\mu_0}-\omega[\![\varepsilon]\!]\right)\left[E_y\right] \end{aligned} \tag{7.43}$$

and transverse magnetic (TM):

$$\begin{aligned} \frac{d}{dz}\left[E_x\right]&=-i\alpha Q_{\varepsilon,zz}^{-1}Q_{\varepsilon,zx}\left[E_x\right]-i\left(\frac{\alpha}{\omega}Q_{\varepsilon,zz}^{-1}\alpha-\omega\mu_0\right)\left[H_y\right] \\ \frac{d}{dz}\left[H_y\right]&=i\omega\left(Q_{\varepsilon,xx}-Q_{\varepsilon,xz}Q_{\varepsilon,zz}^{-1}Q_{\varepsilon,zx}\right)\left[E_x\right]-iQ_{\varepsilon,xz}Q_{\varepsilon,zz}^{-1}\alpha\left[H_y\right] \end{aligned} \tag{7.44}$$

7.4.1.1. Fourier transformation of the permittivity

The set of ordinary differential equations to be integrated contains the Fourier transforms of ε, 1/ε, μ, 1/μ, N_x^2, and N_z^2. In general, Fast Fourier transform (FFT) techniques can be easily applied. As already discussed with respect to eq.(7.39), the normal vector components must be transformed only once, if chosen to be independent on z. On the other hand, the permittivity and permeability depend on z and their Fourier components have to be calculated for each value of z during the numerical integration. Fortunately, in the 1D case, it is possible and recommended to use analytical formulae for the Fourier transforms of ε, 1/ε, μ, 1/μ, which give faster more accurate results. This can be done because for a given value of z, they are piecewise constant functions of y. Fig.7.3 presents schematically a grating with a period d that separates two homogeneous media with permittivities ε_1 and ε_3. For a given value z_0 of z, the Fourier transform of, for example, the permittivity intside the modulated region $0 \le z \le h$ is given by

$$\begin{aligned}\varepsilon_m &= \frac{\varepsilon_1}{d}\int_{x_1}^{x_2} e^{-imK_x x}dx + \frac{\varepsilon_3}{d}\int_{x_2}^{d+x_1} e^{-imK_x x}dx \\ &= (\varepsilon_1-\varepsilon_3)\frac{\sin\left(mK_x\dfrac{x_2-x_1}{2}\right)}{\pi m}e^{-imk_x\frac{x_1+x_2}{2}} + \varepsilon_3\delta_{m,0}\end{aligned} \tag{7.45}$$

so that the two integrals can be solved analytically, once x_1 and x_2 are determined from the inverse of g(x):

$$x_{1,2} = g^{-1}(z_0)\,. \tag{7.46}$$

If the inverse of g(x) has more than two solutions, the sum of integrals (7.45) will contain several more terms. The same equations can be used to obtain the Fourier transforms of the inverse of the permittivity.

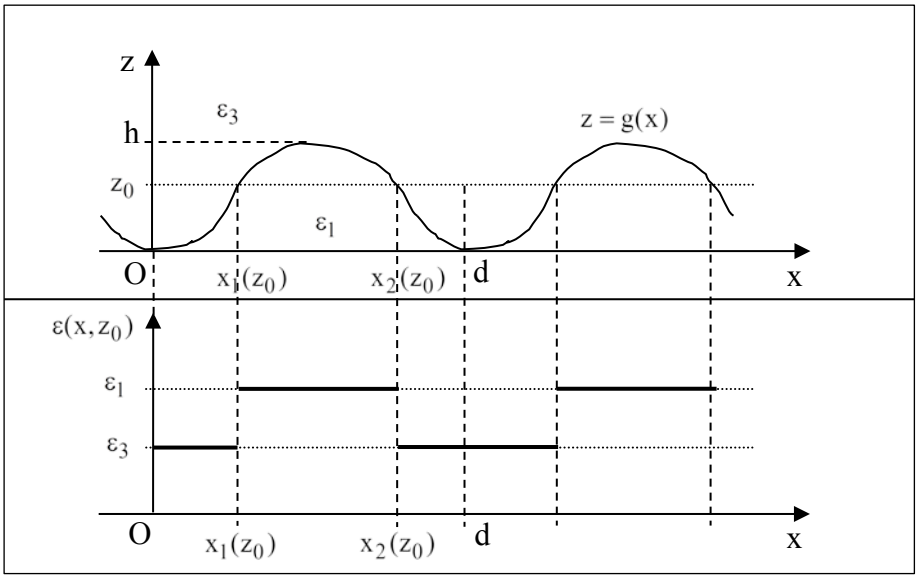


Fig.7.3.Piecewise constant representation of the permittivity for a one-dimensional grating

In the case of a sinusoidal profile:

$$z = \frac{h}{2}\left[1 + \sin(K_x x)\right] \quad \Rightarrow \quad x_{1,2} = \arcsin\left(\frac{2z_0}{h} - 1\right). \tag{7.47}$$

7.4.1.2. Fourier transformation of the normal vector

As already explained, the Fourier transformation of the normal vector requires its continuation all over the space. If the grating profile can be represented as a single-value function, we can use eq.(7.39) for $\vec{N}$ and calculate the Fourier components of the tensor $\vec{N}\vec{N}^T$ by use of the Fast Fourier transform (FFT) technique once for all z-values. For a sinusoidal grating having a profile defined in eq.(7.47), the normal vector takes the form:

$$\vec{N} = \frac{1}{\sqrt{1 + g'^2(x)}}\left(-g'(x), 0, 1\right) = \frac{\left(-\frac{\pi h}{d}\cos(K_x x), 0, 1\right)}{\sqrt{1 + \left(\frac{\pi h}{d}\right)^2 \cos^2(K_x x)}} \tag{7.48}$$

.

7.4.2. Classical isotropic trapezoidal or triangular grating

A trapezoidal grating is shown schematically in Fig.7.4 with two flat regions L at the top and the bottom of the groove and two different, in general, groove angles ψ. The Fourier transform of the permittivity and its inverse are calculated using eq.(7.45) with:

$$\begin{aligned} x_1 &= z_0 \mathrm{cotg}\psi_1 \\ x_2 &= x_C - z_0 \mathrm{cotg}\psi_2 \end{aligned} \tag{7.49}$$

with $x_C = d - L_2$. For the normal vector, the period can be divided in four regions A to D, as shown in the figure:

$$\begin{aligned} &N_y = 0 \\ &\left.\begin{aligned} N_x &= \sin\psi_1 \\ N_z &= -\cos\psi_1 \end{aligned}\right\} \text{ in A,} \qquad \left.\begin{aligned} N_x &= \sin\psi_2 \\ N_z &= \cos\psi_2 \end{aligned}\right\} \text{ in C,} \\ &\left.\begin{aligned} N_x &= 1 \\ N_z &= 0 \end{aligned}\right\} \text{ in B and D,} \end{aligned} \tag{7.50}$$

Their Fourier components do not depend on z and can be represented as a sum of several analytical terms, similar to eq.(7.45):

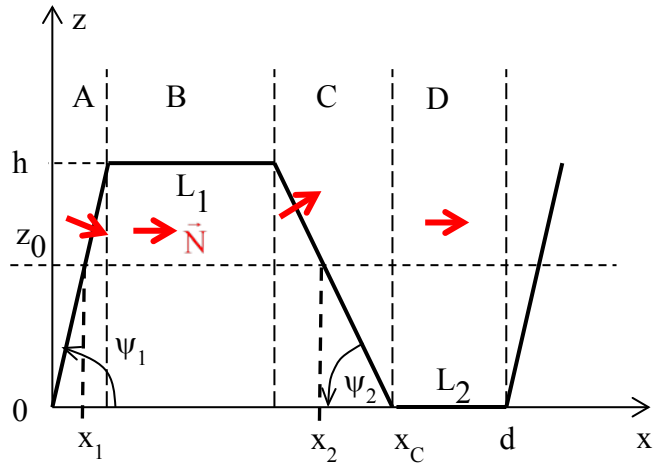


Fig.7.4. Trapezoidal profile with parameters. The normal vector direction is given in red arrows.

$$\left(N_x^2\right)_m = \frac{\sin^2\psi_1}{d}\int_0^{h\,\mathrm{cotg}\psi_1} e^{-imK_x x}dx + \frac{1}{d}\int_{h\,\mathrm{cotg}\psi_1}^{h\,\mathrm{cotg}\psi_1+L_1} e^{-imK_x x}dx$$
$$+\frac{\sin^2\psi_2}{d}\int_{h\,\mathrm{cotg}\psi_1+L_1}^{d-L_2} e^{-imK_x x}dx + \frac{1}{d}\int_{d-L_2}^{d} e^{-imK_x x}dx \qquad (7.51)$$

$$\left(N_y^2\right)_m = \frac{\cos^2\psi_1}{d}\int_0^{h\,\mathrm{cotg}\psi_1} e^{-imK_x x}dx + \frac{\cos^2\psi_2}{d}\int_{h\,\mathrm{cotg}\psi_1+L_1}^{d-L_2} e^{-imK_x x}dx \qquad (7.52)$$

$$\left(N_x N_y\right)_m = -\frac{\sin 2\psi_1}{2d}\int_0^{h\,\mathrm{cotg}\psi_1} e^{-imK_x x}dx + \frac{\sin 2\psi_2}{2d}\int_{h\,\mathrm{cotg}\psi_1+L_1}^{d-L_2} e^{-imK_x x}dx. \qquad (7.53)$$

A triangular-groove grating can be considered as a particular case of a trapezoidal profile with no flat regions, $L_1 = L_2 = 0$, $x_C = d$. Moreover, the profile given in Fig.7.4 also includes the case with vertical facets, and some more exotic profiles with hanging back walls, Fig.7.5.

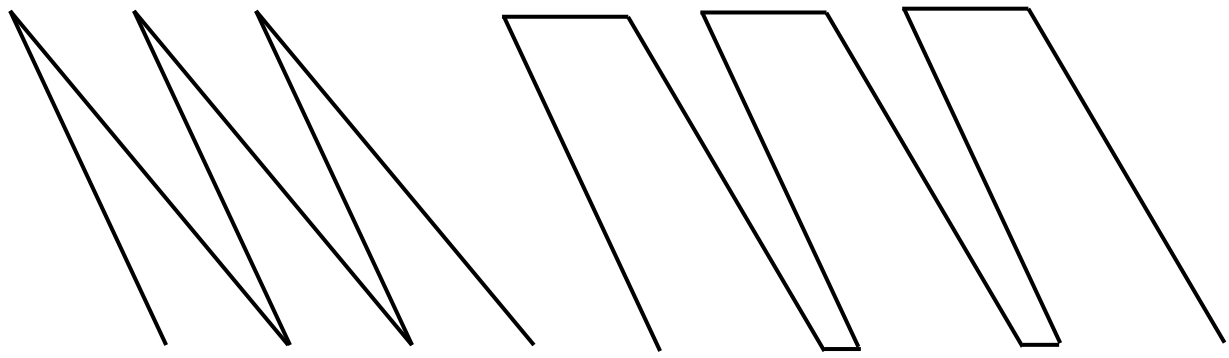


Fig.7.5. Two different profiles with slanted grooves

7.4.3. Classical lamellar grating

Lamellar profile with vertical walls is most easy to treat, because the normal to the profile vector has only one non-zero component, $N_x = 1$. The Q_ε matrix takes the form:

$$Q_\varepsilon = \begin{pmatrix} \left[\!\left[\frac{1}{\varepsilon}\right]\!\right]^{-1} & 0 & 0 \\ 0 & [\![\varepsilon]\!] & 0 \\ 0 & 0 & [\![\varepsilon]\!] \end{pmatrix} \tag{7.54}$$

$$M = \begin{pmatrix} 0 & 0 & \frac{\alpha}{\omega}[\![\varepsilon]\!]^{-1}\beta_0 & -\frac{\alpha}{\omega}[\![\varepsilon]\!]^{-1}\alpha + \omega\mu_0 \mathbb{I} \\ 0 & 0 & \frac{\beta_0^2}{\omega}[\![\varepsilon]\!]^{-1} - \omega\mu_0\mathbb{I} & -\frac{\beta_0}{\omega}[\![\varepsilon]\!]^{-1}\alpha \\ -\frac{\alpha}{\omega\mu_0}\beta_0 & \frac{\alpha^2}{\omega\mu_0} - \omega[\![\varepsilon]\!] & 0 & 0 \\ \omega\left[\!\left[\frac{1}{\varepsilon}\right]\!\right]^{-1} - \frac{\beta_0^2}{\omega\mu_0}\mathbb{I} & \frac{\beta_0}{\omega\mu_0}\alpha & 0 & 0 \end{pmatrix} \tag{7.55}$$

In non-conical diffraction, when $\beta_0 = 0$, the two fundamental polarizations are decoupled and can be solved independently of each other. The M-matrix is simplified to obtain an antidiagonal block form:

$$M = \begin{pmatrix} 0 & 0 & 0 & -\frac{\alpha}{\omega}[\![\varepsilon]\!]^{-1}\alpha + \omega\mu_0\mathbb{I} \\ 0 & 0 & -\omega\mu_0\mathbb{I} & 0 \\ 0 & \frac{\alpha^2}{\omega\mu_0} - \omega[\![\varepsilon]\!] & 0 & 0 \\ \omega\left[\!\left[\frac{1}{\varepsilon}\right]\!\right]^{-1} & 0 & 0 & 0 \end{pmatrix} \tag{7.56}$$

thus the two sets of differential equations for each polarization become:

$$\begin{aligned} \frac{d}{dz}[E_x] &= i\left(\omega\mu_0\mathbb{I} - \frac{\alpha}{\omega}[\![\varepsilon]\!]^{-1}\alpha\right)[H_y] \\ \frac{d}{dz}[H_y] &= i\omega\left[\!\left[\frac{1}{\varepsilon}\right]\!\right]^{-1}[E_x] \end{aligned} \tag{7.57}$$

and

$$\begin{aligned} \frac{d}{dz}[E_y] &= -i\omega\mu_0[H_x] \\ \frac{d}{dz}[H_x] &= i\left(\frac{\alpha^2}{\omega\mu_0} - \omega[\![\varepsilon]\!]\right)[E_y] \end{aligned} \tag{7.58}$$

Even in the case of conical diffraction, it is possible to define two other polarizations, for which the differential system decouples. These are the electric and magnetic polarizations that are *transverse with respect to the x-axis*. Let us denote the two polarization with superscripts (e), when $E_x = 0$, and (h), when $H_x = 0$. For (e) case, it is possible to express H_y as a function of H_x from eq.(7.26) and the first line of the M-matrix in eq.(7.55):

$$\left[H_y^{(e)}\right] = -\left(\omega\mu_0\mathbb{I} - \frac{\alpha}{\omega}[\![\varepsilon]\!]^{-1}\alpha\right)^{-1}\frac{\alpha}{\omega}[\![\varepsilon]\!]^{-1}\beta_0\left[H_x^{(e)}\right] \tag{7.59}$$

which can be simplified into:

$$\left[H_y^{(e)}\right] = -\alpha\left(\omega^2\mu_0[\![\varepsilon]\!] - \alpha^2\right)^{-1}\beta_0\left[H_x^{(e)}\right] \tag{7.60}$$

The second line of the matrix M then results in:

$$\frac{d}{dz}\left[E_y^{(e)}\right] = i\left[\frac{\beta_0^2}{\omega}[\![\varepsilon]\!]^{-1} - \omega\mu_0\mathbb{I} + \frac{\beta_0^2}{\omega}[\![\varepsilon]\!]^{-1}\alpha^2\left(\omega^2\mu_0[\![\varepsilon]\!] - \alpha^2\right)^{-1}\right]\left[H_x^{(e)}\right] \tag{7.61}$$

This expression can be further simplified, and together with the third line of eq.(7.55) (when $E_x = 0$) gives a set of equation for (e) polarization:

$$\begin{aligned}\frac{d}{dz}\left[E_y^{(e)}\right] &= i\omega\mu_0\left[\beta_0^2\left(\omega^2\mu_0[\![\varepsilon]\!] - \alpha^2\right)^{-1} - \mathbb{I}\right]\left[H_x^{(e)}\right]\\ \frac{d}{dz}\left[H_x^{(e)}\right] &= \frac{i}{\omega\mu_0}\left(\alpha^2 - \omega^2\mu_0[\![\varepsilon]\!]\right)\left[E_y^{(e)}\right]\end{aligned} \tag{7.62}$$

Similar procedure for (h) case when $H_x^{(h)} = 0$, result in another system of differential equations, decoupled from the (e) case:

$$\begin{aligned}\frac{d}{dz}\left[H_y^{(h)}\right] &= i\omega\left[\left[\!\!\left[\frac{1}{\varepsilon}\right]\!\!\right]^{-1} + \beta_0^2\left(\alpha[\![\varepsilon]\!]^{-1}\alpha - \omega^2\mu_0\mathbb{I}\right)^{-1}\right]\left[E_x^{(h)}\right]\\ \frac{d}{dz}\left[E_x^{(h)}\right] &= \frac{i}{\omega}\left(\omega^2\mu_0\mathbb{I} - \alpha[\![\varepsilon]\!]^{-1}\alpha\right)\left[H_y^{(h)}\right]\end{aligned} \tag{7.63}$$

In non-conical mount, $\beta_0 = 0$ and eqs.(7.62) and (7.63) become equivalent to eqs.(7.58) and (7.57).

Both conical and nonconical cases of diffraction by lamellar gratings are solved by eigenvector technique, due to the fact that the coefficients of the differential equations are z-independent. Moreover, due to the separation of the two fundamental polarizations, it is possible to further reduce by half the size of the matrices, by dealing with second-order differential equations. For example, eq. (7.57) can be written in the form:

$$\frac{d}{dz}\left[E_x\right] = iM_{14}\left[H_y\right]$$
$$\frac{d}{dz}\left[H_y\right] = iM_{41}\left[E_x\right] \tag{7.64}$$

Thus

$$\frac{d^2}{dz^2}\left[E_x\right] = -M_{14}M_{41}\left[E_x\right]$$
$$\frac{d^2}{dz^2}\left[H_y\right] = -M_{41}M_{14}\left[H_y\right] \tag{7.65}$$

Les us denote with ρ_p^2 the eigenvalues of the product $M_{14}M_{41}$ and with V the matrix with its eigenvectors arranged in columns. The solution of the first eq.(7.65) can be written as:

$$\left[E_x(z)\right] = V\Phi(z)V^{-1}\left[E_x(0)\right] \tag{7.66}$$

with

$$\Phi_{pp'}(z) = \delta_{pp'}\exp(\pm i\rho_p z) \tag{7.67}$$

which shows that the elementary solutions along z (called *modes*, wherefrom the names *Fourier modal method* or *Rigorous coupled waves method*) exist in pairs that can propagate upwards or downwards with the same propagation constants.

By integrating the second eq.(7.65), we obtain that:

$$\left[H_y(z)\right] = W\Phi(z)W^{-1}\left[H_y(0)\right] \tag{7.68}$$

with W that can be written in different forms, because the eigenvectors are defined within an arbitrary factor. For example, if we take into account the second eq.(7.64), $W = \mp i\rho^{-1}M_{41}V$, where the diagonal matrix ρ has elements equal to ρ_p. Another possibility, that is quite convenient in TM polarization described by eq.(7.57), is at first to calculate the eigenvectors of $M_{41}M_{14}$, instead of $M_{14}M_{41}$ (their eigenvalues are the same). Then the link between V and W contains the inverse of M_{41}, which is just equal to $\frac{1}{\omega}\left[\!\left[\frac{1}{\varepsilon}\right]\!\right]$, so that $W = \pm\frac{i}{\omega}\left[\!\left[\frac{1}{\varepsilon}\right]\!\right]V\rho$.

7.4.4. Crossed grating having vertical walls made of isotropic material

Most of the recent applications of the Fourier modal method are devoted to studies of light diffraction by structures with 2D periodicity and piecewise invariant in the third direction. This popularity has several reasons. First, extraordinary light transmission was found in the late 90s by Ebbesen [7.18] for such structures, namely metallic sheets with periodical hole arrays, and it attracted a lot of attention (see Chapter 1). Second, the technology of such structures has significantly advanced in the last 20 years. Third, the Fourier modal method is relatively simple to implement, and much faster than most of the other methods, because of the eigenvalue/vector technique of integration.

Detailed study of these structures will be described in a separate chapter. However, due to its importance, we are discussing different aspects of this theory, as it presents a particular case of the more general geometry, that is characterized by a constant value of the z-component of the normal vector on each cross-section having z = const. The prolongation of the normal vector within the entire grating cell is discussed in sections 7.6.2.2.

7.5. Differential theory for anisotropic media

If we consider anisotropic media that do not extend inside the grating structure, there is not necessary to reformulate the diffraction theory, only that in the general case it is not possible to separate the problem into two independent polarizations, and it is necessary to work with the complete 4Pmax vectors and matrices.

In the case of anisotropic medium that lies inside the grating, the equations linking the M-matrix with the Q_ε and Q_μ matrices remain the same, eq.(7.28) for isotropic and anisotropic media. The difference comes from the fact that the Q-matrices take more complex form, because the link between the normal and tangential components of the couples E and D and H and B is made through the tensors of permittivity and permeability, which are not scalars. Let us establish this link in detail for E and D. As far as the continuous and discontinuous field components must be factorized differently, we construct a column vector F_ε, which contains the continuous field components E_T and D_N. There are two tangential components to the grating surface, and only a single normal:

$$F_\varepsilon = \begin{pmatrix} D_N \\ E_{T_1} \\ E_{T_2} \end{pmatrix} = \begin{pmatrix} \vec{N}\cdot\left(\overline{\overline{\varepsilon}}\,\vec{E}\right) \\ \vec{T}_1\cdot\vec{E} \\ \vec{T}_2\cdot\vec{E} \end{pmatrix} = U_\varepsilon \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} \tag{7.69}$$

where the double bar indicates a second-rank tensor with 3 dimensions, and the matrix U_ε has the form:

$$U_\varepsilon = \begin{pmatrix} \left(\vec{N}\,\overline{\overline{\varepsilon}}\right)_x & \left(\vec{N}\,\overline{\overline{\varepsilon}}\right)_y & \left(\vec{N}\,\overline{\overline{\varepsilon}}\right)_z \\ T_{1x} & T_{1y} & T_{1z} \\ T_{2x} & T_{2y} & T_{2z} \end{pmatrix} \tag{7.70}$$

with $\vec{N}\,\overline{\overline{\varepsilon}}$ representing a tensor product with contraction of indices, for example, $\left(\vec{N}\,\overline{\overline{\varepsilon}}\right)_x = N_x\overline{\overline{\varepsilon}}_{xx} + N_y\overline{\overline{\varepsilon}}_{yx} + N_z\overline{\overline{\varepsilon}}_{zx}$, etc.

The vectors $\vec{N}, \vec{T}_1, \text{and}\, \vec{T}_2$ are defined on the grating surface, but for their further Fourier transform, it is necessary to choose a suitable continuation. The necessary conditions are that (i) they are continuous on the surfaces where ε and μ are discontinuous, and (ii) they form an orthonormal triad.

Since $\overline{\overline{\varepsilon}}$ never vanishes, the determinant of U_ε represents a quadric non-null form, equal to:

$$\xi_\varepsilon \equiv \det U_\varepsilon = \left(\vec{N}\,\overline{\overline{\varepsilon}}\right)\cdot\left(\vec{T}_1\times\vec{T}_2\right) = \sum_{i,j=x,y,z} N_i\varepsilon_{ij}N_j \tag{7.71}$$

since $\vec{N} = \vec{T}_1\times\vec{T}_2$.

Thus U_ε has an inverse U_ε^{inv} in the form:

$$U_\varepsilon^{inv} = \frac{1}{\xi_\varepsilon}\begin{pmatrix} N_x & -\left[\left(\vec{N}\,\overline{\overline{\varepsilon}}\right)\times\vec{T}_2\right]_x & \left[\left(\vec{N}\,\overline{\overline{\varepsilon}}\right)\times\vec{T}_1\right]_x \\ N_y & -\left[\left(\vec{N}\,\overline{\overline{\varepsilon}}\right)\times\vec{T}_2\right]_y & \left[\left(\vec{N}\,\overline{\overline{\varepsilon}}\right)\times\vec{T}_1\right]_y \\ N_z & -\left[\left(\vec{N}\,\overline{\overline{\varepsilon}}\right)\times\vec{T}_2\right]_z & \left[\left(\vec{N}\,\overline{\overline{\varepsilon}}\right)\times\vec{T}_1\right]_z \end{pmatrix} \tag{7.72}$$

It is not evident to derive this form, but it can easily be verified by using the equivalence $U_\varepsilon U_\varepsilon^{inv} = \mathbb{I}$ and the fact that $\vec{N} = \vec{T}_1 \times \vec{T}_2$. For example, the product of the second line of U_ε with the second column of U_ε^{inv} can be written in vectorial form:

$$\xi_\varepsilon\left(U_\varepsilon U_\varepsilon^{inv}\right)_{yy} = -\vec{T}_1\cdot\left[\left(\vec{N}\,\overline{\overline{\varepsilon}}\right)\times\vec{T}_2\right] = -\left[\left(\vec{N}\,\overline{\overline{\varepsilon}}\right)\times\vec{T}_2\right]\cdot\vec{T}_1 = -\left(\vec{N}\,\overline{\overline{\varepsilon}}\right)\cdot\left(\vec{T}_2\times\vec{T}_1\right) = \left(\vec{N}\,\overline{\overline{\varepsilon}}\right)\cdot\vec{N} = \xi_\varepsilon \tag{7.73}$$

Going back to the vector F_ε, it is continuous across the grating surface, whereas the Cartesian components of the electric vector are, in general discontinuous, as well as the components of U_ε. Thus for their Fourier transform, it is necessary to apply the inverse factorization rule:

$$\left[F_\varepsilon\right] = \left[\!\left[U_\varepsilon^{inv}\right]\!\right]^{-1}\left[\vec{E}\right] \tag{7.74}$$

At the other hand,

$$\vec{E} = U_\varepsilon^{inv}F_\varepsilon \quad\Rightarrow\quad \vec{D} = \overline{\overline{\varepsilon}}\,U_\varepsilon^{inv}F_\varepsilon \tag{7.75}$$

with F_ε being continuous, so that the Fourier transform of $\vec{D}$ requires the direct factorization rule:

$$\left[\vec{D}\right] = \left[\!\left[\overline{\overline{\varepsilon}}\,U_\varepsilon^{inv}\right]\!\right]\left[\!\left[U_\varepsilon^{inv}\right]\!\right]^{-1}\left[\vec{E}\right] \tag{7.76}$$

i.e.,

$$Q_\varepsilon = \left[\!\left[\overline{\overline{\varepsilon}}\,U_\varepsilon^{inv}\right]\!\right]\left[\!\left[U_\varepsilon^{inv}\right]\!\right]^{-1} \tag{7.77}$$

For gratings having anisotropic magnetic properties, the corresponding Q_μ matrix is obtained from eqs. (7.71), (7.72), and (7.77) by replacing U_ε^{inv} by U_μ^{inv} and $\overline{\overline{\varepsilon}}$ by $\overline{\overline{\mu}}$.

7.5.1. Lamellar gratings made of anisotropic material

Such gratings are analyzed in the chapter devoted to the Fourir modal method by using more direct approaches, but here we want show how the corresponding equations can be obtained from the general eqs.(7.72). To this aim it is sufficient to realize that

$$\begin{aligned} \vec{N} &= (1,0,0) \\ \vec{T}_1 &= (0,1,0) \\ \vec{T}_2 &= (0,0,1) \end{aligned} \tag{7.78}$$

so that

$$\begin{aligned}\left(\vec{N}\,\overline{\overline{\varepsilon}}\right)\times\vec{T}_2 &= (\varepsilon_{xy}, -\varepsilon_{xx}, 0)\\ \left(\vec{N}\,\overline{\overline{\varepsilon}}\right)\times\vec{T}_1 &= (-\varepsilon_{xz}, 0, \varepsilon_{xx})\\ \left(\vec{N}\,\overline{\overline{\varepsilon}}\right)\cdot\vec{N} &= \varepsilon_{xx}\end{aligned} \tag{7.79}$$

and eq.(7.72) takes the form:

$$U_{\varepsilon}^{inv} = \begin{pmatrix} \dfrac{1}{\varepsilon_{xx}} & -\dfrac{\varepsilon_{xy}}{\varepsilon_{xx}} & -\dfrac{\varepsilon_{xz}}{\varepsilon_{xx}} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{7.80}$$

with a determinant equal to $\left[\!\left[\dfrac{1}{\varepsilon_{xx}}\right]\!\right]$. Thus

$$\left[\!\left[U_{\varepsilon}^{inv}\right]\!\right]^{-1} = \begin{pmatrix} \left[\!\left[\dfrac{1}{\varepsilon_{xx}}\right]\!\right]^{-1} & \left[\!\left[\dfrac{1}{\varepsilon_{xx}}\right]\!\right]^{-1}\left[\!\left[\dfrac{\varepsilon_{xy}}{\varepsilon_{xx}}\right]\!\right] & \left[\!\left[\dfrac{1}{\varepsilon_{xx}}\right]\!\right]^{-1}\left[\!\left[\dfrac{\varepsilon_{xz}}{\varepsilon_{xx}}\right]\!\right] \\ 0 & \mathbb{I} & 0 \\ 0 & 0 & \mathbb{I} \end{pmatrix} \tag{7.81}$$

The second matrix that is required takes the form:

$$\overline{\overline{\varepsilon}}U_{\varepsilon}^{inv} = \begin{pmatrix} 1 & 0 & 0 \\ \dfrac{\varepsilon_{yx}}{\varepsilon_{xx}} & \varepsilon_{yy} - \dfrac{\varepsilon_{yx}\varepsilon_{xy}}{\varepsilon_{xx}} & \varepsilon_{yz} - \dfrac{\varepsilon_{yx}\varepsilon_{xz}}{\varepsilon_{xx}} \\ \dfrac{\varepsilon_{zx}}{\varepsilon_{xx}} & \varepsilon_{zy} - \dfrac{\varepsilon_{zx}\varepsilon_{xy}}{\varepsilon_{xx}} & \varepsilon_{zz} - \dfrac{\varepsilon_{zx}\varepsilon_{xz}}{\varepsilon_{xx}} \end{pmatrix} \tag{7.82}$$

This form is valid even when the permittivity tensor is not symmetric, as happens in the modeling of magnetooptical effects.

The Q_ε matrix takes the form obtained in [7.19], using a completely different approach:

$$Q_\varepsilon = \left[\!\left[\bar{\bar{\varepsilon}}\, U_\varepsilon^{inv} \right]\!\right] \left[\!\left[U_\varepsilon^{inv} \right]\!\right]^{-1} \tag{7.83}$$

$$= \begin{pmatrix} \left[\!\left[\frac{1}{\varepsilon_{xx}}\right]\!\right]^{-1} & \left[\!\left[\frac{1}{\varepsilon_{xx}}\right]\!\right]^{-1}\left[\!\left[\frac{\varepsilon_{xy}}{\varepsilon_{xx}}\right]\!\right] & \left[\!\left[\frac{1}{\varepsilon_{xx}}\right]\!\right]^{-1}\left[\!\left[\frac{\varepsilon_{xz}}{\varepsilon_{xx}}\right]\!\right] \\ \left[\!\left[\frac{\varepsilon_{yx}}{\varepsilon_{xx}}\right]\!\right]\left[\!\left[\frac{1}{\varepsilon_{xx}}\right]\!\right]^{-1} & \left[\!\left[\varepsilon_{yy} - \frac{\varepsilon_{yx}\varepsilon_{xy}}{\varepsilon_{xx}}\right]\!\right] + \left[\!\left[\frac{\varepsilon_{yx}}{\varepsilon_{xx}}\right]\!\right]\left[\!\left[\frac{1}{\varepsilon_{xx}}\right]\!\right]^{-1}\left[\!\left[\frac{\varepsilon_{xy}}{\varepsilon_{xx}}\right]\!\right] & \left[\!\left[\varepsilon_{yz} - \frac{\varepsilon_{yx}\varepsilon_{xz}}{\varepsilon_{xx}}\right]\!\right] + \left[\!\left[\frac{\varepsilon_{yx}}{\varepsilon_{xx}}\right]\!\right]\left[\!\left[\frac{1}{\varepsilon_{xx}}\right]\!\right]^{-1}\left[\!\left[\frac{\varepsilon_{xz}}{\varepsilon_{xx}}\right]\!\right] \\ \left[\!\left[\frac{\varepsilon_{zx}}{\varepsilon_{xx}}\right]\!\right]\left[\!\left[\frac{1}{\varepsilon_{xx}}\right]\!\right]^{-1} & \left[\!\left[\varepsilon_{zy} - \frac{\varepsilon_{zx}\varepsilon_{xy}}{\varepsilon_{xx}}\right]\!\right] + \left[\!\left[\frac{\varepsilon_{zx}}{\varepsilon_{xx}}\right]\!\right]\left[\!\left[\frac{1}{\varepsilon_{xx}}\right]\!\right]^{-1}\left[\!\left[\frac{\varepsilon_{xy}}{\varepsilon_{xx}}\right]\!\right] & \left[\!\left[\varepsilon_{zz} - \frac{\varepsilon_{zx}\varepsilon_{xz}}{\varepsilon_{xx}}\right]\!\right] + \left[\!\left[\frac{\varepsilon_{zx}}{\varepsilon_{xx}}\right]\!\right]\left[\!\left[\frac{1}{\varepsilon_{xx}}\right]\!\right]^{-1}\left[\!\left[\frac{\varepsilon_{xz}}{\varepsilon_{xx}}\right]\!\right] \end{pmatrix}$$

7.6. Normal vector prolongation for 2D periodicity; Fourier transform

As observed, the proper use of the direct and the inverse factorization rules requires that the vector normal to the interfaces between different media is defined not only on these interfaces, but throughout the entire grating cell. In the case of classical gratings with one-dimensional periodicity, the prolongation of the normal vector can be done quite easily, as shown in sec.7.4.1. For two-dimensional periodicity, the choice depends on the geometry, but also on its mathematical representation. Several different solutions have to be considered, without pretending to be exhaustive.

In general, the cross-section profile changes with z, so that the matrices Q_ε, Q_μ, and M have to be recalculated for each value of z. If the geometry is z-invariant, this must be done only once. Concerning the Fourier components of the normal vector, there are two different classes of grating profiles that has to be treated separately. The first class consists of surfaces that can be expressed all over the unit cell (containing a single period in x and z) as an analytical (at lease piecewisely) function $z_S = g(x, y)$, where S indicates that the point lies on the interface. In this case, it is possible to have a unique extension of $\vec{N}$ whatever the values of z. In addition, it is not necessary to calculate the cross-section of the surface(s) with a plane perpendicular to the z-axis for each value of z. This case also includes multilayered homomorphous structure with constant layer(s) thickness in the z-direction. We consider this class of cases in sec.7.6.1.

The second class of surfaces includes surfaces that cannot be expressed through single-valued functions, as the example given on the right-hand side of Fig.7.12. In that case, it is necessary for each fixed value of z to know the cross-section function of the grating surface with the plane z = const. Subsection 7.6.2 presents general analysis, some important specific cases are considered further in the following subsections.

7.6.1. General analytical surfaces

If the interface representing the structure can be expressed as a single-valued function, analytical over the entire unit cell (this is also valid if different analytical functions can be defined over different regions of the cell):

$$z_S = g(x, y), \tag{7.84}$$

then the components of the vector normal to the surface defined on the surface have the form:

$$\vec{N}(x,y)=\frac{\left(-\frac{\partial g}{\partial x},-\frac{\partial g}{\partial y},1\right)}{\sqrt{1+\mathrm{grad}^2 g(x,y)}}. \tag{7.85}$$

It can immediately be extended to whatever the values of z inside the modulated region. Moreover, its values do not depend on z, as it was a case of the classical one-dimensional grating already discussed in sec.7.4.1.

The permittivity and its inverse can easily be obtained on a mesh (x,y) covering the grating cell for each z:

$$\begin{aligned} g(x,y)<z \quad &\Rightarrow \quad \varepsilon(x,y)=\varepsilon_3 \\ g(x,y)\geq z \quad &\Rightarrow \quad \varepsilon(x,y)=\varepsilon_1 \end{aligned} \tag{7.86}$$

where 1 and 3 are the indexes representing respectively the inferior and the superior regions separated by the grating surface (as it was done in Fig.7.3). If the cross-section of the grating surface with the planes z = const are ellipses (or circles), and if N_z does not depend on (x,y) at each z, it is possible to replace the numerical Fourier transform by an analytical formulae. One important particular case is the z-invariant grating with elliptical cross section; another case includes the gratings having a rotational symmetry, as shown in Fig.7.12.

The same extension (7.85) for the normal vector is valid for a stack of layers having homogeneous thicknesses in the z-direction:

$$z_{S,j}=g(x,y)+t_j \tag{7.87}$$

The permittivity and its inverse inside the intermediate layers are simply given as:

$$z_{S,j-1}(x,y)<z\leq z_{S,j}(x,y) \quad \Rightarrow \quad \varepsilon(x,y)=\varepsilon_j. \tag{7.88}$$

7.6.2. Irregular general surfaces

If the case does not fit into the preceding section, the interface is expressed through the more general function $u(x,y,z_S)=0$, and the vector $\vec{N}$ has to be determined for each inclusion:

$$\vec{N}(x,y,z_S)=\frac{\left(\frac{\partial u}{\partial x},\frac{\partial u}{\partial y}\frac{\partial u}{\partial z_S}\right)}{\left|\mathrm{grad}\, u(x,y,z_S)\right|} \tag{7.89}$$

However, these values are well defined on the grating surface (except on its edges), and have to be extended over the entire cell. When considerintg a cross-section of the profile with a plane at z = const., several different cases exist:

7.6.2.1. Single-valued radial cross-section

At first, we shall consider that the cross section function $f(x_S,y_S)=0$ defines a single curve, and that curve can be expressed in cylindrical coordinates as

$$\rho_S = \rho_S(\varphi) \tag{7.90}$$

where

$$\begin{aligned} \rho_S &= \sqrt{(x_S - x_C)^2 + (y_S - y_C)^2} \\ \varphi &= \arctan\left[(y_S - y_C)/(x_S - x_C)\right] \end{aligned} \tag{7.91}$$

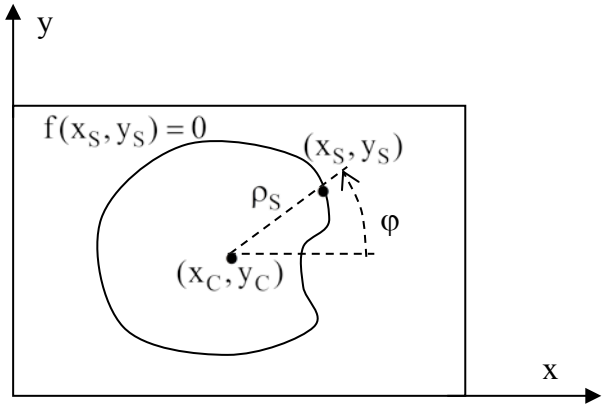


Fig.7.6. Single-curve cross-section of the grating surface at z = const.

and x_C and y_C represent a "central" point of the curve, Fig.7.6. Here we assume that the values of ρ_S are unique for each φ. The other case is analyzed further on.

It is possible to extend to the entire cell the values of the normal vector, defined only on the curve, by assuming that it is constant for each fixed angle φ. This prolongation requires the following procedure:

1. Fixing the pair (x,y).
2. Calculating the angle $\varphi = \arctan\left[(y - y_C)/(x - x_C)\right]$.
3. Calculating $\rho_S = \rho_S(\varphi)$ from eq.(7.91).
4. Calculation of $x_S = \rho_S \cos\varphi$, and $y_S = \rho_S \sin\varphi$.
5. Determining N_z, together with N_x and N_y from eq.(7.89).
6. Attributing these values of the components of $\vec{N}(x_S, y_S)$ to the pair (x,y).
7. Fast Fourier transform after the normal vector components are determined for all the pairs (x,y) on a mesh inside the grating cell.

The procedure can be simplified for most of the typical diffracting objects, as shown further for objects with elliptical or circular cross-section.

If the grating profile varies with z, the calculations of the Fourier components of the permittivity and its inverse has to be made at each value of z, both for the analytical profiles, for which the normal vector prolongation can be chosen z-invariant, or for the irregular surfaces. For each (x,y) pair of the mesh used in the FFT method, it is possible to determine whether the point lies within or outside the cross-section part of Fig.7.6:

$$\begin{aligned} &\rho(\varphi) < \rho_S(\varphi), \quad \varepsilon(x,y) = \varepsilon_{\text{inside}} \\ &\rho(\varphi) \geq \rho_S(\varphi), \quad \varepsilon(x,y) = \varepsilon_{\text{outside}} \end{aligned} \tag{7.92}$$

with $\rho = \sqrt{(x - x_C)^2 + (y - y_C)^2}$.

7.6.2.2. Objects with polygonal cross section

A typical example of such objects is presented in Fig.7.1. Its surface consists of different plates, and for their treatment the condition N_z = const. for fixed z is fulfilled, because $\vec{N}$ is constant at each plate. Another possible surface consists of plane ribbons with curvature in z-direction, Fig.7.7.

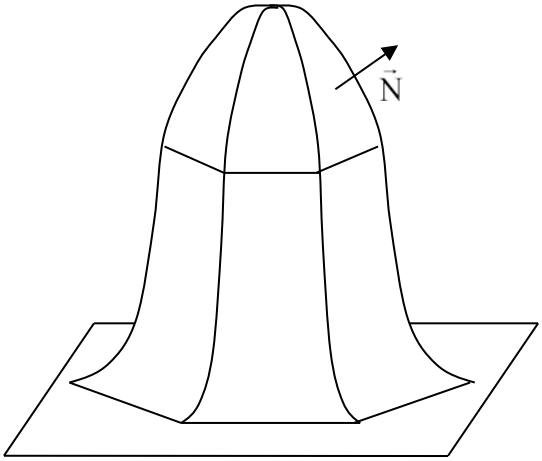


Fig.7.7. Surface made of plane ribbons

As shown in Fig.7.8, the cross-section represents a polygon. On each of its sides the modulus of the in-plane component of the normal vector is known:

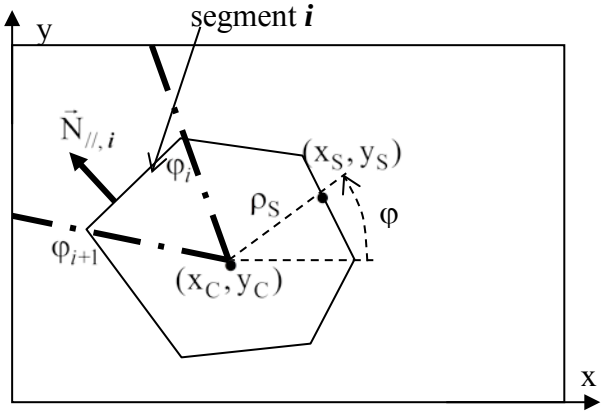


Fig.7.8. Object with a polygonal cross-section.

$$N_{//,i} = \sqrt{1 - N_{z,i}^2} \tag{7.93}$$

and its direction is perpendicular to the segment. If the *i*-th segment is located between the angles φ_i and φ_{i+1}, we can extend the definition of the normal vector all over the unit cell situated within the range (φ_i, φ_{i+1}), delimited by the bold dot-dashed lines in Fig.7.8, by

assuming that $\vec{N} = \vec{N}_i$. The normal vector extension will be continuous everywhere, except on the sector border lines (bold dot-dashed lines) and thus the only points where both permittivity (and/or permeability) and $\vec{N}$ are simultaneously discontinuous are at the polygonal corners, where anyway $\vec{N}$ is never continuous.

The procedure to follow requires that for each value of z the polygon corners (x_i, y_i) and the z-components of $\vec{N}$ for each segment are determined, as well as fixing some "central" point (x_C, y_C). Then the angular ranges of each segment with respect to that central point are calculated:

$$\varphi_i = \arctan\frac{y_i - y_C}{x_i - x_C} \tag{7.94}$$

For each pair (x,y), the azimuthal angle is given as $\varphi = \arctan\left[(y - y_C)/(x - x_C)\right]$, which value determines the number of the segment, say j, within the point lies. The unknown in-plane part of the normal vector is perpendicular to the j-th segment:

$$\begin{aligned} N_{x,j} &= \left(y_{j+1} - y_j\right)\sqrt{\frac{1 - N_{z,j}^2}{\left(y_{j+1} - y_j\right)^2 + \left(x_{j+1} - x_j\right)^2}} \\ N_{y,j} &= -\left(x_{j+1} - x_j\right)\sqrt{\frac{1 - N_{z,j}^2}{\left(y_{j+1} - y_j\right)^2 + \left(x_{j+1} - x_j\right)^2}} \end{aligned} \tag{7.95}$$

The expression in the square root comes from the normalization of $\vec{N}$.

The value of the permittivity depends on whether the point (x,y) lies inside or outside the polygon. The calculations of $\varepsilon(x,y)$ and $1/\varepsilon$ are made simultaneously with the normal vector calculus. After the angular segment in which the point (x,y) of the mesh in grating cell is determined (say the j-th one, as in eq.(7.95)), we can find the length of ρ_S between the central point and the polygon segment, shown in Fig.7.8. For this sake we show in Fig.7.9 the enlarged segment:

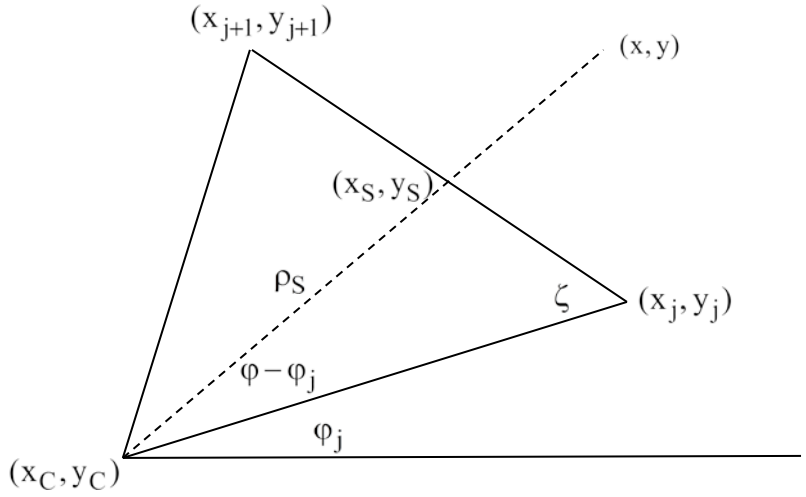


Fig.7.9. The j-segment of Fig.7.8 together with notations

The sine theorem gives the possibility to determine the angle ζ:

$$\frac{\sqrt{\left(x_{j+1}-x_j\right)^2+\left(y_{j+1}-y_j\right)^2}}{\sin(\varphi_{j+1}-\varphi_j)}=\frac{\sqrt{\left(x_{j+1}-x_C\right)^2+\left(y_{j+1}-y_C\right)^2}}{\sin(\zeta)} \tag{7.96}$$

wherefrom the radius ρ_S is given as:

$$\rho_S=\sin(\zeta)\frac{\sqrt{\left(x_C-x_j\right)^2+\left(y_C-y_j\right)^2}}{\sin(\pi-\zeta-\varphi+\varphi_j)} \tag{7.97}$$

Eq.(7.92) enables us to obtain the values of the permittivity (and its inverse):

$$\begin{aligned}&\rho(\varphi)<\rho_S(\varphi), \quad \varepsilon(x,y)=\varepsilon_{inside}\\ &\rho(\varphi)\geq\rho_S(\varphi), \quad \varepsilon(x,y)=\varepsilon_{outside}\end{aligned} \tag{7.98}$$

with $\rho=\sqrt{(x-x_C)^2+(y-y_C)^2}$.

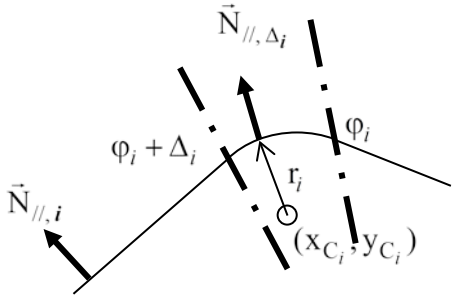


Fig.7.10. Schematic presentation of corner rounding

Concerning the edges, in reality the surfaces never have such, as etching always ends by rounding the corners, as shown in Fig.7.10. Let us consider that the rounding between the segments numbered *i*-1 and *i* is made preserving the values of N_z, and that in the cross-plane z = const., the rounding can be considered as circular, having a center in (x_{C_i},y_{C_i}) and radius r_i .The in-plane component of the normal vector follows the curvature radius and thus is given by equations, similar to eq.(7.95):

$$\begin{aligned}N_{x,\Delta_i}&=\left(x-x_{C_i}\right)\sqrt{\frac{1-N_{z,i}^2}{\left(y-y_{C_i}\right)^2+\left(x-x_{C_i}\right)^2}}\\ N_{y,\Delta_i}&=\left(y-y_{C_i}\right)\sqrt{\frac{1-N_{z,i}^2}{\left(y-y_{C_i}\right)^2+\left(x-x_{C_i}\right)^2}}\end{aligned} \tag{7.99}$$

The prolongation is more complicated, if the consecutive values of N_z at the two sides of the rounded corner differ significantly. In that case a linear interpolation of N_z between φ_i and $\varphi_i + \Delta_i$ can be applied.

7.6.2.3. Mutlivalued cross-sections

If the cross-section cannot be represented as a radial function, nother possibility arises if it is a piecewise analytical function in x (or y), as shown in Fig.7.11, where we can use two different functions of x. We assume again that N_z is known, as it happens for z-independent profiles, for which it is simply null. In the upper part of the figure, for each value of x it is possible to calculate the normal vector on the profile:

$$\vec{N}_1 = \frac{\left(-f_1'(x), 1, N_z\right)}{\sqrt{1 + f_1'^2(x) + N_z^2}} \tag{7.100}$$

We can take this value to be the same for each y in the upper region A_1, so that the numerical Fourier transform is made only once in A_1 and once in A_2.

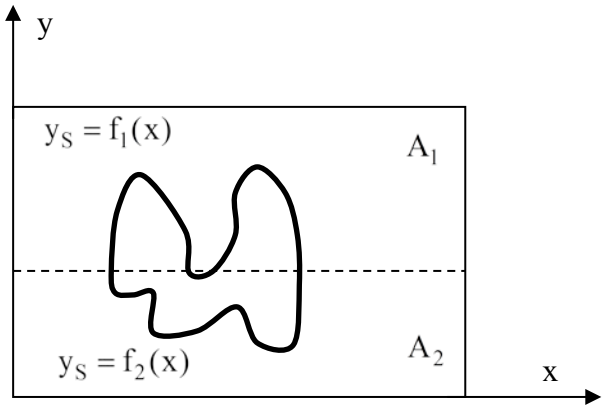


Fig.7.11. Piecewise analytical c ross-section of the grating surface at z = const.

The permittivity has to be calculated at each (x,y) mesh point:

$$\begin{aligned} y < y_S, \quad & \varepsilon(x,y) = \varepsilon_{\text{inside}} \\ y \geq y_S, \quad & \varepsilon(x,y) = \varepsilon_{\text{outside}}. \end{aligned} \tag{7.101}$$

7.6.4. Objects with cylindrical symmetry

Many periodic systems consist of inclusion having rotational cylindrical symmetry, like spheres, vertical cylinders, or ellipsoids with axe of rotation parallel to the z-axis, but also smooth surfaces, as presented in Fig.7.12.

These structures are characterized by a circular cross-section of the surface with the horizontal planes at z = const., but also with an independence on x and y of the values of N_z on each horizontal plane. In addition, due to the circular cross-sections, the angular component $N_\varphi = 0$ everywhere. Once z is fixed, the variation of the interface in the vertical

direction fixes the value of N_z, for example through eq.(7.89), wherefrom the radial normal vector component $N_\rho = \sqrt{1 - N_z^2}$. For each pair of (x,y) then:

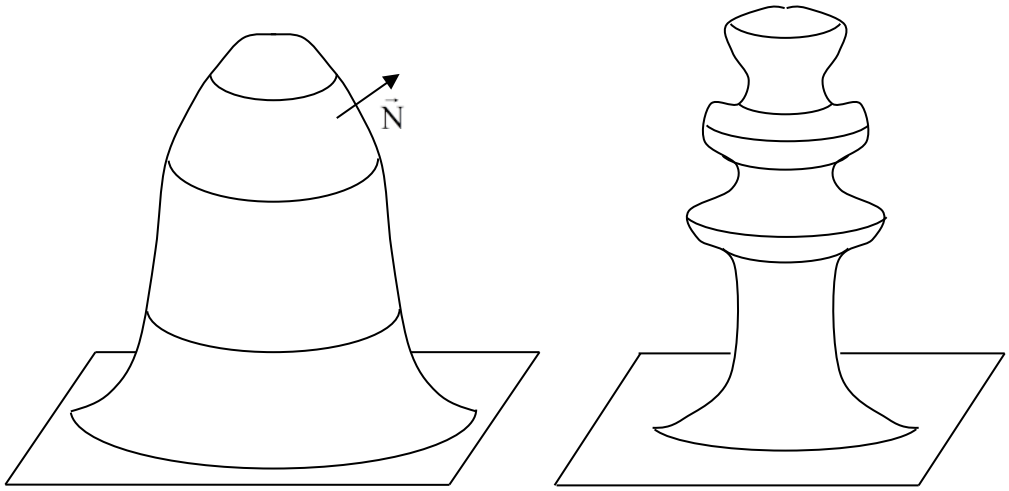


Fig.7.12. Several profiles with cylindrical symmetry.

$$N_x(x,y) = \frac{(x - x_C)N_\rho}{\sqrt{(x - x_C)^2 + (y - y_C)^2}}$$
$$N_y(x,y) = \frac{(y - y_C)N_\rho}{\sqrt{(x - x_C)^2 + (y - y_C)^2}} \qquad (7.102)$$

In addition, for profiles invariant in z-direction, $N_z = 0$ and $N_\rho = 1$.

The permittivity is given as a piecewise constant function:

$$\varepsilon(x,y) = \varepsilon_{inside}, \quad \text{if} \quad (x - x_C)^2 + (y - y_C)^2 < R^2(z),$$
$$\varepsilon(x,y) = \varepsilon_{outside}, \quad \text{if} \quad (x - x_C)^2 + (y - y_C)^2 \geq R^2(z) \qquad (7.103)$$

where R(z) is the radius of the profile surface for a given z.

Having obtained the values of the normal vector components and permittivity for each x,y enables us to calculate their Fourier transforms, either by Fast Fourier transform (FFT), or analytically.

7.6.5. Objects with elliptical cross-section

Similar simplification is possible for systems with elliptical cross-sections that have N_z = const. for z fixed. Such are the inclusions of vertical cylinders with elliptical cross-section, ellipsoids with one of the axes orientated in z-direction, but also all types of the structures shown schematically in Fig.7.12 that have elliptical or circular cross-sections.

Let us assume that the ellipse axes are parallel to the x and y-axes.. The cross-section curve for z = const. is given by the equation:

$$\left(\frac{x_S - x_C}{a}\right)^2 + \left(\frac{y_S - y_C}{b}\right)^2 = R^2 \tag{7.104}$$

In order to obtain results similar to eq.(7.102), we introduce an elliptical coordinates, defined as:

$$\begin{aligned} \tilde{x} &= \frac{x - x_C}{a} \\ \tilde{y} &= \frac{y - y_C}{b} \end{aligned} \tag{7.105}$$

Using these notations, the ellipse becomes a circle, for which the considerations of the previous subsection apply. Thus

$$\begin{aligned} N_x(x,y) &= \frac{\frac{x - x_C}{a} N_\rho}{\sqrt{\left(\frac{x - x_C}{a}\right)^2 + \left(\frac{y - y_C}{b}\right)^2}} \\ N_y(x,y) &= \frac{\frac{y - y_C}{b} N_\rho}{\sqrt{\left(\frac{x - x_C}{a}\right)^2 + \left(\frac{y - y_C}{b}\right)^2}} \end{aligned} \tag{7.106}$$

with $N_\rho = \sqrt{1 - N_z^2}$, which remains constant for each fixed z.

Concerning the permittivity, it is determined in the same way as in eq.(7.103) for circular profile:

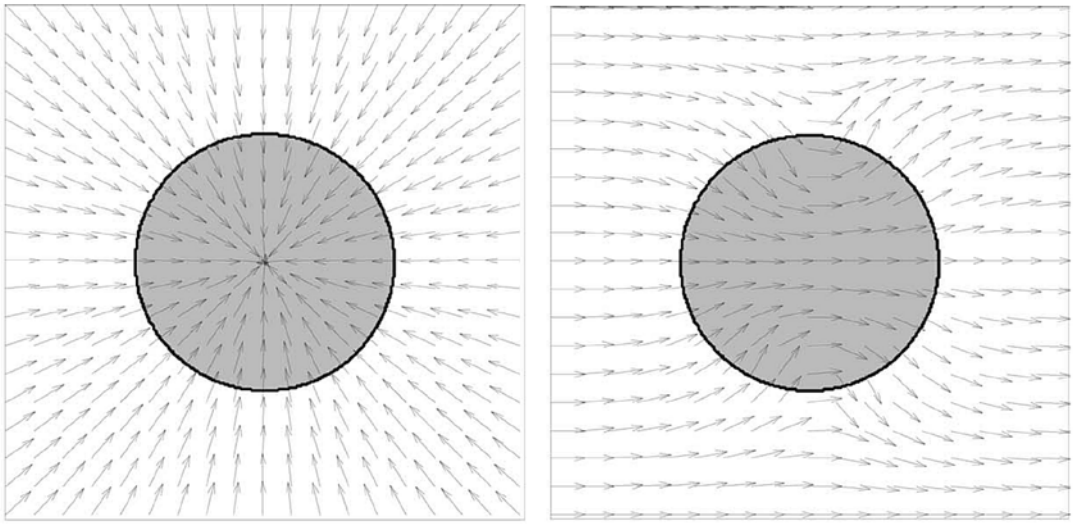
$$\begin{aligned} &\varepsilon(x,y) = \varepsilon_{inside}, \quad \text{if} \quad \left(\frac{x - x_C}{a}\right)^2 + \left(\frac{y - y_C}{b}\right)^2 < R^2(z), \\ &\varepsilon(x,y) = \varepsilon_{outside}, \quad \text{if} \quad \left(\frac{x - x_C}{a}\right)^2 + \left(\frac{y - y_C}{b}\right)^2 \geq R^2(z) \end{aligned} \tag{7.107}$$

with R(z) given by eq.(7.104).

Remark on the prolongation of the normal vector

Special attention has recently been paid to the numerical implementation of the differential method for gratings having 2D periodicity formed by vertical holes or bumps that are invariant in z, and that have arbitrary cross-section in the xOy plane. A detailed study in the case of z-invariant geometry that applies for an eigenvalue/eigenvector technique of integration (FM or RCW method) is given in ref.[7.20], followed by several other works [7.21, 7.22]. It is necessary to note that the technique of prolongation of the normal vector as discussed in [7.20] can be applied also for z-dependent profiles with similar cross section; the difference is the renormalization factor $\sqrt{1 - N_z^2}$ for each z.

The authors compare several different formulations of the Fourier Modal method applied to structures with rectangular, circular, or elliptical cross-sections. These formulations include the classical formulation of Moharam and Gaylor that uses only the direct factorization rule, the formulation given by Lifeng Li [7.23] that introduces two different Fourier transforms of the permittivity ε, namely $\lfloor\lceil\varepsilon\rceil\rfloor$ and $\lceil\lfloor\varepsilon\rfloor\rceil$, which are calculated by applying at first the inverse rule along one of the coordinates, and then the direct rule along the other one. This second formulation was made for rectangular and parallelogram cross-sections. For circular or elliptical (or other smooth) forms, it introduces a stepwise treatment of the profile, which appears more slowly convergent than the special techniques developed after.



(a) (b)

Fig.7.13. Two different prolongations of the normal vector for a circular inclusion. (a) Radial prolongation. (b) Electrostatic continuation of the normal vector for a circular cross-section inclusion inside a square grating cell (after [7.20]).

The third approach to the problem requires a prolongation of the normal vector (NV) to the profile within the entire grating cell. As already stressed, there are several possibilities to make this. A typical example is the radial prolongation, Fig.7.13a, which has been discussed in Sec.7.6.2.1 and 7.6.4 and it includes discontinuities of the normal vector on the cell boundaries, where the permittivity is continuous. Another approach proposed in [7.20] is the electrostatic one, which insures the continuity all over the cell and on its boundaries, except for on single points inside, Fig.7.13b.

Fig.7.14 shows the convergence rates for the transmitted zeroth order of a grating consisting of dielectric cylindrical inclusions with a circular cross-section with refractive index n = 1.5, in normal incidence from the substrate. The grating period is 2λ, the width of the grooves is λ, and the grating depth is $\lambda/(2n-1)$. The graph presents the diffraction efficiency in transmission as a function of the truncation order N using the three considered formulations: Moharam's original formulation, Li's formulation, and the formulation using the normal vector (NV) field. As usual the Fourier series run from – N to N, which yields 2N+1 Fourier coefficients for each of the two directions of periodic continuation, or $(2N+1)^2$ coefficients in total. As can be expected, both the original approach and the formulation by Li have worse convergence than when correctly taking into account the factorization rules for the tangential electric field and normal displacement components to the profile, where the permittivity is discontinuous [7.20]. It is necessary to stress that the difference in the

convergence rates is even more pronounced for metallic inclusions, having much larger optical contrast.

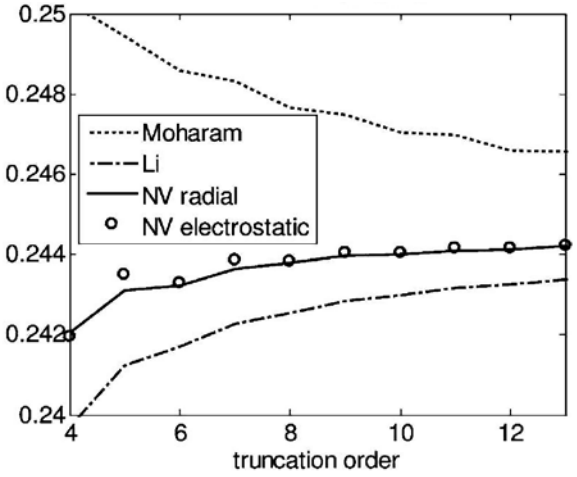
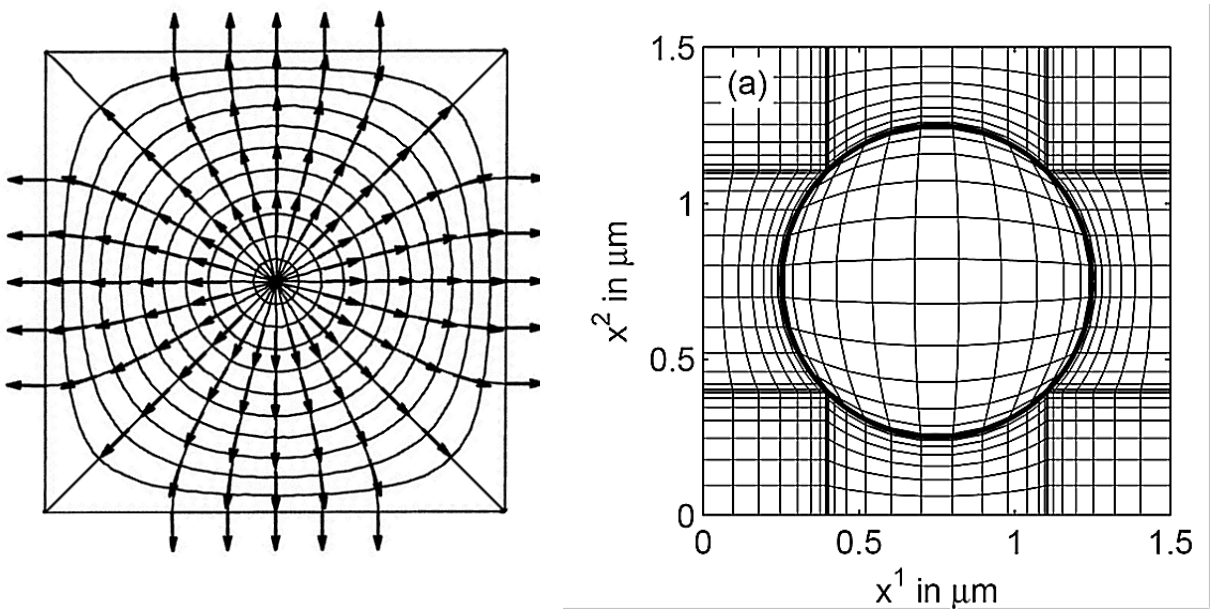


Fig.7.14. Convergence rates with respect to the truncation of the Fourier series for four different approaches used to model the diffraction by a cylindrical inclusion with circular cross-section.

Recently, Weiss et al. [7.24] proposed another alternative approach to treating smooth inclusions, by changing the coordinate system, so that its planes are parallel to the profile *and* to the grating sell walls (see Fig.7.15). If the transformed system is orthonormal, its coordinate lines are automatically tangential or perpendicular to the physical walls. If not, the Maxwell equations have to be rewritten in covariant vector form using the covariant and contravariant vector components.



(a) (b)

Fig.7.15. Coordinate lines and surfaces according to (a) [7.20] and (b) [7.24]

This approach is somehow equivalent to the normal vector prolongation, due to two main reasons:

(i) The NV approach defines in an unambiguous manner the normal vectors on the profile, giving a liberty to continue them all over the cell. The coordinate transformation is also defined on the profile and the outside boundaries, but can be chosen in different ways around the grating cell.

(ii) The change of the coordinate system introduces in the Maxwell equations the metric tensor $\mathbb{G}$ that multiplies the electric displacement and magnetic induction in the right-hand side of eq.(7.6), so that for the electric field we obtain the substitution:

$$\vec{D} = \varepsilon\vec{E} \quad \rightarrow \quad \mathbb{G}\varepsilon\vec{E}. \tag{7.108}$$

The normal vector approach acts in a similar manner by introducing the matrix Q_ε, given in eq.7.21, which makes the following substitution in the Fourier space, eq.(7.20):

$$\left[\vec{D}\right] = \left[\varepsilon\vec{E}\right] \quad \rightarrow \quad Q_\varepsilon\left[\vec{E}\right]. \tag{7.109}$$

7.6.6. Multiprofile surfaces

A grating with multiple bumps (or inclusions) inside the single cell could be treated by separation the cell into sub-cells, not necessarily rectangular, containing a single inclusion, as shown in Fig.7.16, where a specific cross-section at z = const. is separated into three regions A, B, and C. As far as the Fourier components of the normal vector, of the permeability and the permittivity have to be calculated for each value of z (if they depend on z), the separation into subcells can vary with z.

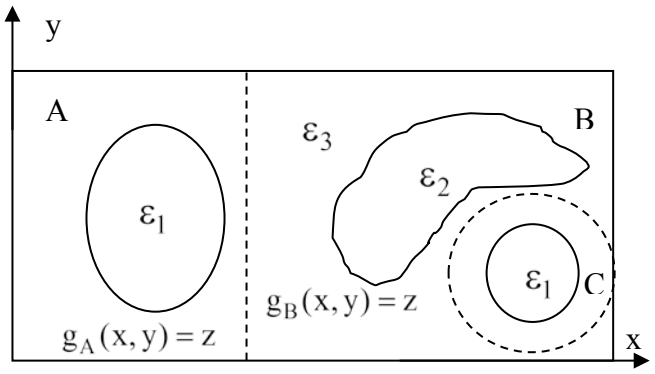


Fig.7.16. Cross-section at z = const. of a grating having different inclusions. The three different regions to be treated independently are separated by dashed lines.

The case schematized in Fig.7.17a can result from a surface covered by a thin layer of another substance, a layer that cannot be treated using eq.(7.87). The simplest possibility is to have different continuation of the normal vector inside each region. At first, the angle $\varphi = \arctan\left[(y - y_C)/(x - x_C)\right]$ for the point with coordinates (x, y) is calculated, and it is necessary to determine to which region the point belong. If it lies inside the innermost region C, the values of $\vec{N}(x,y) = \vec{N}_1(\varphi)$, where $\vec{N}_1(\varphi)$ is determined using one of the procedures discussed above for a single interface that is defined by the inner profile function.

If the point (x, y) lies in the outermost region, we take $\vec{N}(x,y) = \vec{N}_2(\varphi)$, where $\vec{N}_2(\varphi)$ corresponds to the second interface. In-between, we have two possibilities. The first choice is to divide the region into two subregions as indicated in Fig.7.17a with the dashed line. In each of them, $\vec{N}(x,y)$ is taken to be equal to its values on the adjacent profile, so that it is continuous everywhere where the permittivity and/or permeability are discontinuous.

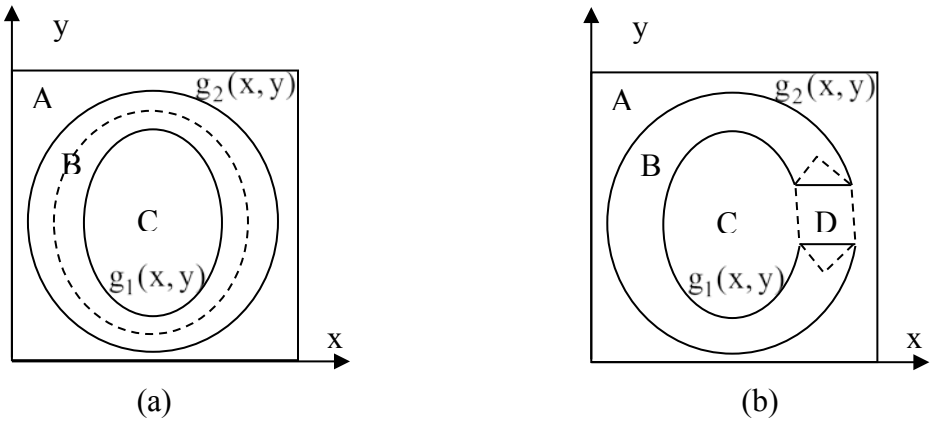


Fig.7.17. Structures with interpenetrating cross-section profiles

The second possibility is to introduce a linear interpolation inside region B, but it is necessary to know the distances ρ_1 and ρ_2 between the central point and the profiles along the ray with $\varphi = \arctan\left[(y - y_C)/(x - x_C)\right]$ fixed. Then:

$$\vec{N}_B(x,y) = \vec{N}_1(\varphi) + \left[\vec{N}_2(\varphi) - \vec{N}_1(\varphi)\right]\frac{\rho - \rho_1}{\rho_2 - \rho_1}, \quad \text{with} \quad \rho^2 = (x - x_C)^2 + (y - y_C)^2 \tag{7.110}$$

Another specific case that appears in the studies of magnetic resonators is presented in Fig.7.17b. In can be treated in the same way is for the case in Fig.7.17a, but it is necessary to introduce a separate region D indicated in the figure and containing the opening, for which $\vec{N} = (0, 1, 0)$, for example.

7.7. Integrating schemes

Numerical solution of a system of ordinary differential equations is a mature domain due to the enormous amount of physical and technical applications. Unfortunately, the grating problem represents one of the worst tasks for the theory of ordinary differential equations, because the system to be integrated is a *stiff* one. To better understand the problem, let us consider the case of a homogeneous layer that introduces no coupling between the diffraction orders. The solution of the diffraction problem contains waves propagating up- and downwards (in z-direction). These are plane waves, propagating or evanescent inside the layer. In lossless medium, their constant of propagation in z can be real, or imaginary, depending on the number of diffraction order under consideration:

$$\gamma_m = \pm\sqrt{(k_0 n)^2 - \alpha_m^2}\,, \tag{7.111}$$

with

$$\alpha_m = \alpha_0 + mK\,. \tag{7.112}$$

The real values of γ are bounded by $k_0 n$, but the imaginary parts are not bounded, as their asymptotic values for large |m| are given by:

$$\mathrm{Im}(\gamma_m) = \pm |m| K\,. \tag{7.113}$$

From the point of view of the theory of ordinary differential equations this means that the eigenvalues of the system differ significantly in magnitude, i.e., the differential system is *stiff*. The greater the difference, the more unstable the solution. On the other hand, the solution of the diffraction problem requires sufficient number of Fourier components of the profile function and electromagnetic field to be correctly represented by the truncated Fourier series, thus the necessity to work with large number of Fourier components, and thus the increasingly greater the stiffness of the differential system, i.e., more instable the solution with respect to the length and number of integration steps. The theory concludes that the so called *explicit* integration schemes are most instable for such problem, whatever their order, and *implicit* methods have to be used. The problem with the implicit methods is that they need one matrix inversion and several more matrix operations on each integration step, when compared with the explicit methods, so that the choice is not evident to ensure the most efficient integration scheme.

Let us recall the basic principle of the first-order explicit and implicit schemes. In a first-order approximation, the solution of the differential system:

$$\frac{d}{dz}F(z) = M(z)F(z). \tag{7.114}$$

between two consecutive points $z = z_j$ and $z = z_{j+1}$ can be searched in developing in series:

$$F(z_{j+1}) = F(z_j) + (z_{j+1} - z_j)M(z)F(z). \tag{7.115}$$

If $M(z)$ and $F(z)$ are evaluated in $z = z_j$, this leads to the first-order explicit integration (Euler's) scheme:

$$F(z_{j+1}) = \left[\mathbb{I} + hM(z_j)\right]F(z_j). \tag{7.116}$$

where $\mathbb{I}$ is the unit matrix, and $h = (z_{j+1} - z_j)$.

If $M(z)$ and $F(z)$ are evaluated in $z = z_{j+1}$, we obtain the first-order implicit (inverted or backward Euler's) scheme:

$$F(z_{j+1}) = \left[\mathbb{I} - hM(z_{j+1})\right]^{-1} F(z_j). \tag{7.117}$$

The theory says that this scheme is more stable, but it needs one matrix inversion on each step. A combination of the two must provide even better results, because it uses a half of the previous step:

$$F(z_{j+1}) = \left[\mathbb{I} - \frac{1}{2}hM(z_{j+1})\right]^{-1}\left[\mathbb{I} + \frac{1}{2}hM(z_j)\right]F(z_j). \tag{7.118}$$

However, we need one additional matrix multiplication. In what follows we use these two single-point first-order methods under the names **Expl 1** (single point explicit Euler integration) and **Impl 1**, eq.(7.118) and compare the convergence with respect to the total number of integration points with several other more sophisticated integration schemes for two different metallic gratings in TM polarization, the most difficult combination when using the differential method.

The advantage of these formulations is that they all are single-step ones, and do not need a storage of the intermediate results on several integration steps. They can be easily programmed and don't need additional memory storage at each step. However, if we refer to one of the most relevant sources [7.25], we see that "*this is the generic disease of stiff equations: We are required to follow the variation in the solution on the shortest length scale to maintain the stability of the integration, even though accuracy requirements allow a much larger stepsize*." This means that *a priori* choice of the integration step without adaptive control and change in the step length cannot produce stable and relevant results. Unfortunately, it is quite difficult to use adaptive-step methods, because they require much

longer computation times, as it is necessary to repeat the integration process several times when changing the integration step length. This is why we concentrate our attention to fixed-step algorithms.

Fixed-step *multistep explicit* methods have been used from decades in the differential method programming. The best results have been obtained when combined with an implicit correction by using a predictor-corrector scheme, as described further on. However, referring again to [7.25], "*high order* does not always mean *high accuracy*." It will be more useful, if larger integration step is obtained with high order or multistep methods, which is not obvious, as we observe on several numerical examples.

We have used three simple integration schemes, the single-point implicit or explicit scheme, as well as a 4-point predictor-corrector method (**PCM 4**). It contains two steps, the first one representing an Adams-Bashforth explicit 4-point scheme [7.26], described by the equation

$$\begin{aligned} &F(z_{j+5}) = F(z_{j+4}) \\ &+ h\left[\frac{1901}{720}F'(z_{j+4}) - \frac{1387}{360}F'(z_{j+3}) + \frac{108}{30}F'(z_{j+2}) - \frac{637}{360}F'(z_{j+1}) + \frac{251}{720}F'(z_j)\right]. \end{aligned} \tag{7.119}$$

with

$$F'(z_j) = M(z_j)F(z_j). \tag{7.120}$$

The corrector step is a 4-point Adams-Moulton integration:

$$\begin{aligned} &F(z_{j+4}) = F(z_{j+3}) \\ &+ h\left[\frac{251}{720}F'(z_{j+4}) + \frac{646}{720}F'(z_{j+3}) - \frac{264}{720}F'(z_{j+2}) + \frac{106}{360}F'(z_{j+1}) - \frac{19}{720}F'(z_j)\right], \end{aligned} \tag{7.121}$$

which is an implicit scheme [7.27]. However, contrary to the other explicit schemes (BDF), it does not require inverting a matrix, it just makes one step back as a corrector.

An extension of eq.(7.117) to a multistep algorithm results in multistep implicit method, called also backward differentiation formulae (BDF) [7.28]. Typical 3-point and 5-point formulae take the form:

$$\text{BDF3:} \qquad F(z_{j+3}) = \left[\mathbb{I} - hM(z_{j+3})\right]^{-1}\left[\frac{18}{11}F(z_{j+2}) - \frac{9}{11}F(z_{j+1}) + \frac{2}{11}F(z_j)\right]. \tag{7.122}$$

$$\begin{aligned} &\text{BDF5:} \qquad F(z_{j+5}) \\ &= \left[\mathbb{I} - hM(z_{j+5})\right]^{-1}\left[\frac{300}{137}F(z_{j+4}) - \frac{300}{137}F(z_{j+3}) + \frac{200}{137}F(z_{j+2}) - \frac{75}{137}F(z_{j+1}) + \frac{12}{137}F(z_j)\right]. \end{aligned} \tag{7.123}$$

It is evident that BDF3 (called further on **Impl 3**) requires a starting method for the first two points, and BDF5 (**Impl 5**), for the first 4 points. The same is valid for PCM in eqs.(7.119) – (7.121).

The second-order Runge-Kutta method is given in the form:

$$\begin{aligned} k_1 &= hM(z_j)F(z_j) \\ k_2 &= hM(z_{j+1/2})\left[F(z_j) + \frac{1}{2}k_1\right] \\ F(z_{j+1}) &= F(z_j) + k_2 \end{aligned} \tag{7.124}$$

which has an error proportional to h^3. It is also called a midpoint point, because it requires the evaluation of the functions at the middle of the step, i.e., twice the number of the steps of the other tree methods discussed above.

The classical fourth-order Runge-Kutta method also uses midpoint values:

$$
\begin{aligned}
k_1 &= hM(z_j)F(z_j) \\
k_2 &= hM(z_{j+1/2})\left[F(z_j)+\frac{1}{2}k_1\right] \\
k_3 &= hM(z_{j+1/2})\left[F(z_j)+\frac{1}{2}k_2\right] \\
k_4 &= hM(z_{j+1/2})\left[F(z_j)+k_3\right] \\
F(z_{j+1}) &= F(z_j)+\frac{1}{6}k_1+\frac{1}{3}k_2+\frac{1}{3}k_3+\frac{1}{6}k_4
\end{aligned}
\tag{7.125}
$$

with an error of the order of h^5. These two methods are higher-order explicit methods that do not need matrix inversions during the integration. However, as already stressed, higher order does not mean larger steps.

The highest possible order in h can be obtained theoretically, by using the **eigentechnique**:

$$F(z_{j+1}) = V\left(e^{\gamma h}\right)V^{-1}F(z_j) \tag{7.126}$$

where V is a matrix containing the eigenvectors of M and the exponential term in the round brackets represents a diagonal matrix constructed using the eigenvalues γ of M. In practice, this method does not increase the stability, because it remains a single-point explicit method. Moreover, it requires much longer computation time because of the requirement to solve an eigenvalue/eigenvector problem at each integration step.

The first example concerns a typical commercial sinusoidal aluminum grating that has very high efficiency in TM polarization. It supports a single diffracted order in -1^{st} order Littrow mount and has a modulation depth-to-period ratio of 40%. Fig.7.20 presents a numeric test of the efficiency calculated for a different number of integration points using several integrations schemes. Due to the polarization and the grating material, it is necessary to separate the integration into several slices (5 in this case) in order to avoid numerical loss of precision, the results of the integration in two consecutive slices connected to each other by the use of the S-matrix algorithm. In Fig.7.20b we have presented a part of the results, obtained with 20 slices, instead of 5. The comparison between the two cases show that 5 slices are sufficient, the weaker oscillation for 20 slices are due mainly to the fact that the horizontal scale is less dense, because the step in the total number of integration points is an integer times the number of slices. The truncation parameter N = 20, i.e., totally 41 Fourier harmonics of the field are used in the calculations.

As can be concluded, an absolute precision within 1% is rapidly obtained whatever the method used, with the total number of points of the order of 200. However, the predictor corrector method is less stable when the number of points is smaller than 300. Implicit methods are more stable, as expected, and result in an error smaller than 0.1% even for the number of points less than 200. It is interesting to observe than the first-order implicit method is more stable than the higher implicit methods, probably because it contains a middle-point evaluation of the field derivative, as seen in eq.(7.118). It requires a little bit longer computation time than the other two implicit methods, because of the additional matrix multiplication. The explicit method, which is the fastest one, shows slower convergence, as expected, whereas the performance of the higher-order RK methods competes with the implicit methods.

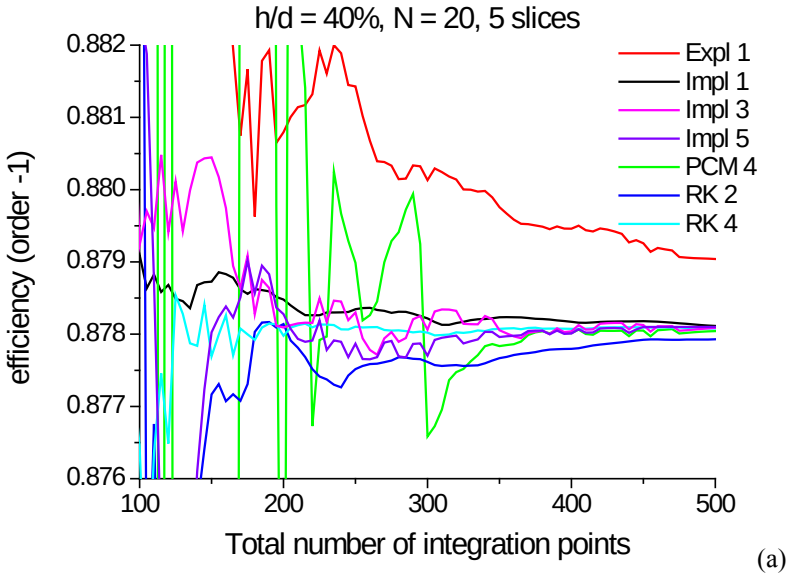


Fig.7.18. Aluminum grating with period 0.5 μm and depth 0.2 μm used in -1st order Littrow mount at 0.6328 μm wavelength in TM polarization. Convergence with respect to the total number of integration points, truncation to 41 Fourier harmonics and using 5 (a) and 20 (b) slices in the S-matrix algorithm. The acronyms for the methods are defined in the text.

Table 7.1 compares the computation times of the different methods for the two investigated cases (Fig.7.18 and the following Fig.7.19). For comparison, (null) indicates the time without any operation due to the integration, and that is necessary for the construction of the M-matrix and the use of the S-matrix propagation algorithm, as described in Appendix 7.A. The fastest method is the single-point explicit method, but as expected it is less precise given the same number of integration points, Fig.7.18. The implicit single-step middle-point method shows stability similar to the 4-th order Runge-Kutta method, but is slightly more rapid. The predictor-corrector method is less stable and requires longer computation times.

Table 7.1. Computation times of the different methods described in the text for the two cases with groove depth to groove period equal to 40% and 200%

Method	N = 20, slices 5 int. points 400, modulation 40%	N = 50, slices 35 int.points 1500, modulation 200%
Expl. 1	1.41 s	72.3 s
Impl.1	3.55 s	187.6 s
Impl.3	2.67 s	157.8 s
Impl.5	2.81 s	168.9 s
PCM 4	2.14 s	120.0 s
RK 2	2.70 s	144.9 s
RK 4	4.70 s	252.5 s
eigentechnique	7.54 s	360.0 s
(null)	0.68 s	38.5 s

It is necessary to stress out that in reality, the computation times are shorter than listed in the Table, because when the truncation N is smaller (usually 20 is sufficient), the number of slices for the S-matrix algorithm is smaller (due to the smaller number of evanescent orders taken into account); in addition, the total number of integration points used for constructing the Table are chosen to obtain 0.1% relative error, whereas in most of the cases just 1% is sufficient. The computation time grows linearly with the total number of integration points, as well as with the number of slices used in the S-matrix algorithm. The time dependence concerning the truncation N in the Fourier series grows as $N^3 - N^{3.5}$, because this parameter determines the size of the matrices.

When the total integration length is multiplied by 5, the number of integration points required is also multiplied by the same factor, as observed in Fig.7.19. A grating twice as deep as the period, acts almost like a flat mirror in TM polarization, with the efficiency in order -1 hardly exceeding 1%. Due to the large depth, the absorption is increased, so that the reflectivity in order 0 is equal to 56.78%. We compare the convergence in the weak -1st order, so that even a small absolute error appears as a large relative error that can be easily observed in the figures. The number of Fourier harmonics (truncation parameter 2N+1) also has to be increased by a factor of 2.5 to 101 (N = 50). The number of slices in the S-matrix algorithm is increased seventh-fold to 35.

The first Fig.7.19 compares several explicit integration schemes with the single-point implicit method. The main conclusion to be drawn is that the best scheme remains the implicit method, only the explicit Runge-Kutta fourth-order scheme seems to compete in convergence rate with respect to the total number of integration points, but somehow slower.

The comparison of several implicit methods confirms the general idea that multistep choice does not necessarily improve the stability (Fig.7.20). When compared with Fig.7.19, the implicit methods are characterized by smaller oscillations when the number of steps is increased, but the most rapid convergence is obtained with the simplest procedure, single-step method (let us remark again that we use the middle-point calculations, as in eq.(7.118)). Like all the other implicit methods, it requires a single matrix inversion on each integration step, but needs less memory storage, and avoids several matrix sums and multiplication by different constants, necessary for the multipoint methods.

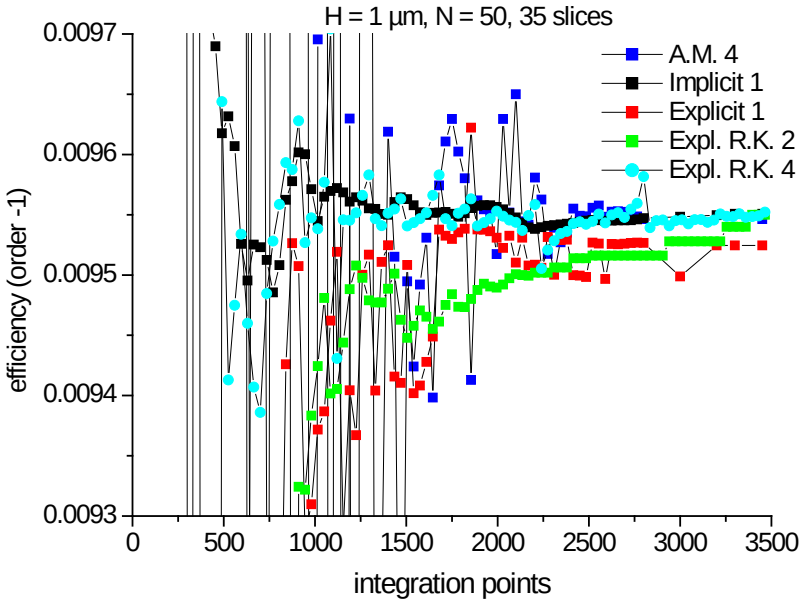


Fig.7.19. Aluminum grating with period 0.5 μm and depth 1 μm used in -1st order Littrow mount at 0.6328 μm wavelength in TM polarization. Convergence with respect to the total number of integration points, truncation to 101 Fourier harmonics and using 35 slices in the S-matrix algorithm.A.M.4 – forth-order Adams-Moulton scheme, Implicit 1 – single-point implicit scheme, Explicit 1 – single-point explicit scheme, Expl2.R.K.2 and 4 – explicit Runge-Kutta method of order 2 and4, respectively.

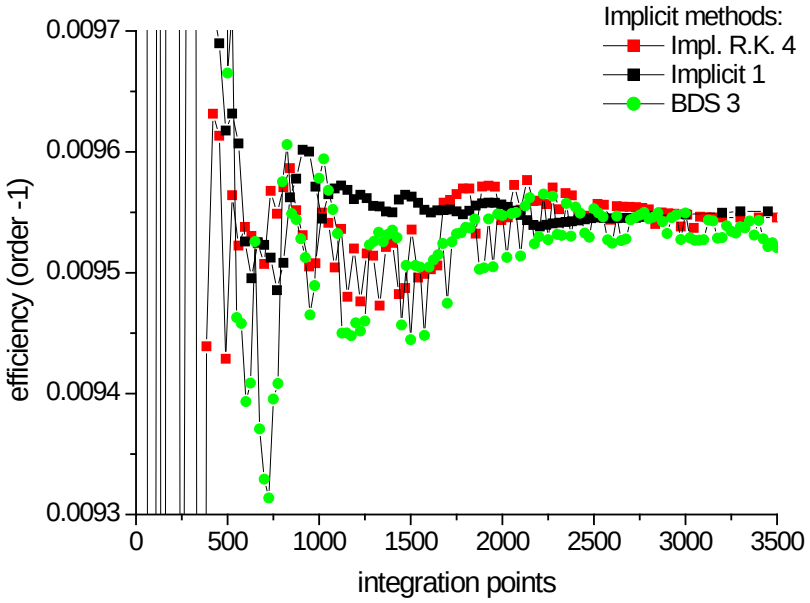


Fig.7.20. Same as in Fig.7.19 but for three implicit methods of different order.

7.8. Staircase approximation

As already discussed, if the surface interface is z-invariant (entirely or piecewisely), the integration of the system of ordinary differential equations along z can be done via eigenvalue/eigenvector technique, eq.(7.126), because the M-matrix containing the coefficients of the differential equations does not depend on z. The enormous interest in this approach can be explained by the simple technique of integration, much easier to understand

and apply than the theory of numerical methods of integration of ordinary differential equations.

The idea is sketched in Fig.7.21, where a sinusoidal surface-relief grating is approximated by a 5-stairs profile. While this approximation (with sufficient number of steps M, depending on the groove depth) works quite well in TE polarization, the TM case presents a convergence rate with respect the truncation number of Fourier components of the field much slower than the ordinary differential method (no staircase approximation), see Fig.7.22. Moreover, the greater the number of vertical slices M, the greater the truncation number required.

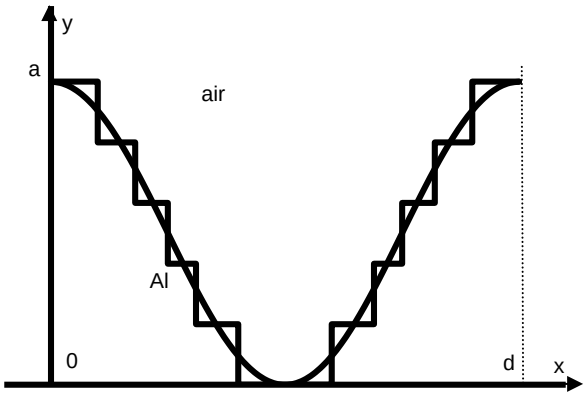


Fig.7.21. Schematical approximation of a sinusoidal grating profile, approximated by a 5-step staircase profile.(after [7.29]).

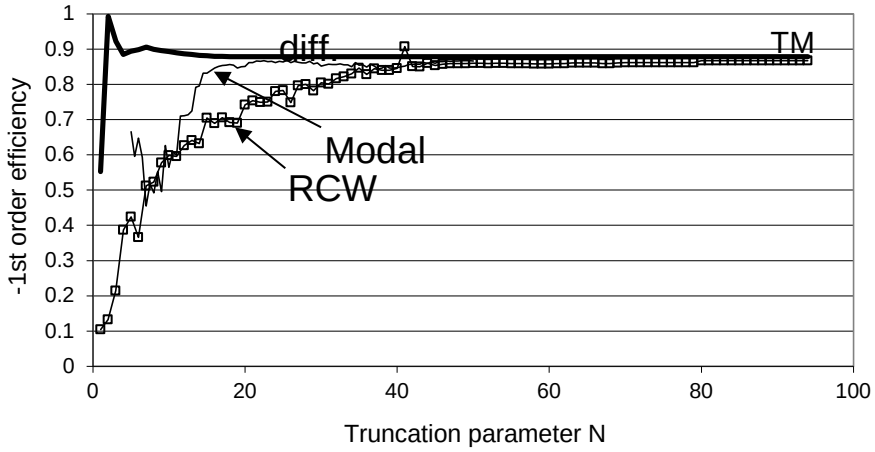


Fig.7.22. Convergence of the minus-first-order efficiency in TM polarization of the FMM (RCW) and the exact modal method (indicated on the figure) for a sinusoidal grating in a staircase presentation with $\tilde{M} = 20$*, as compared to the convergence of the differential method for a smooth sinusoidal profile (curve "diff."). Period d = 0.5 μm, groove depth a = 0.2 μm, aluminum refractive index* $n_{Al} = 1.3 + i7.6$*, illuminated at 40° incidence with wavelength* $\lambda = 0.6328$ *μm, (after [7.29)].*

A detailed analysis of this problem can be found in [7.14, 7.29], but the basic idea is quite simple. The staircase approximation substitutes the otherwise smooth sinusoidal profile by a profile that has sharp edges. The greater the number of stairs, the greater is the number of edges. It is well-known from general electromagnetism that edges introduce electric field singularities. While in TE polarization the only electric field components are tangential to the profile (in y-direction), thus have no discontinuities and singularities, this is not the case in

TM polarization. This can be observed in Fig.7.23. At the edges of each step, a sharp maximum of the electric field is observed. These maxima are not a numerical artifact, they represent the physical effect of introducing edges to replace a smooth profile. These sharp variations of the field require larger number of Fourier components to be correctly represented. Moreover, the greater the number of slices (stairs), the greater the number of the maxima, thus the greater the truncation number required. Numerical experiment has shown that this phenomenon has nothing to do with the integration (eigenvalue/vector) technique, because the results of the convergence rate and field maps are the same for the staircase approximation when using the RCW technique or the differential method.

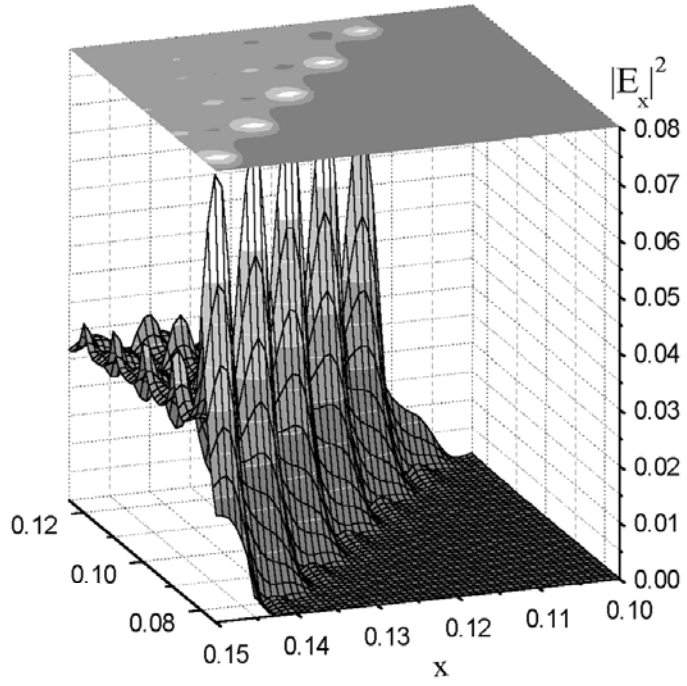


Fig.7.23. Spatial field distribution of $|E_x|^2$ in the vicinity of several steps inside a groove of a 10-step staircase profile, used to approximate the sinusoidal grating under study in TM polarization. The grating parameters are the same as in Fig.7.22, after [7.29].

On the contrary, if the true smooth profile is treated by the differential method by using a numerical integration of the ordinary differential system with the elements of the M-matrix depending on z, there is no such singularities of the electric field (Fig.7.24), so that the convergence with respect to the number of Fourier harmonics is much faster, provide the correct factorization rules are used (Fig.7.22).

Recently, some authors [7.30] have proposed to maintain the eigenvalue/vector technique, but to use the correctly determined Fourier presentation of the profile, i.e., the correct factorization rules, as presented in eqs. (7.39) – (7.44), instead of lamellar-profile factorization, eqs.(7.56) – (7.65), at each step. This is equivalent to using the formulation proposed by the differential method for a smooth profile, i.e. avoiding the field singularities at the edges, but to use the eigenvalue/vector technique of integration by assuming that the modified M-matrix, as given by eqs. (7.39) — (7.44), is z-invariant across each step height. We have already tried this in [7.29] and the conclusion was that using this approach, the number of steps (stairs) has to be relatively larger that by using some better adapted integration technique. And indeed, the eigenvalue/approach to a z-dependent system is equivalent to the rectangular rule with equidistant points of integration, one of the worst choices, as known from the theory of ordinary differential equations. In addition, due to

eigenvalue/vector evaluation on each integration step, its computation times are several times longer than for the other methods (see Table 7.1 in the previous section), known from the theory of ordinary differential equations. This is why the authors of [7.30] need more than 2000 equidistant points of integration for a trapezoidal profile, for which the better adapted numerical integration scheme can suffice with 300 points. Unfortunately, the authors of [7.30] do not consider the differential method as a "reference method" in their work.

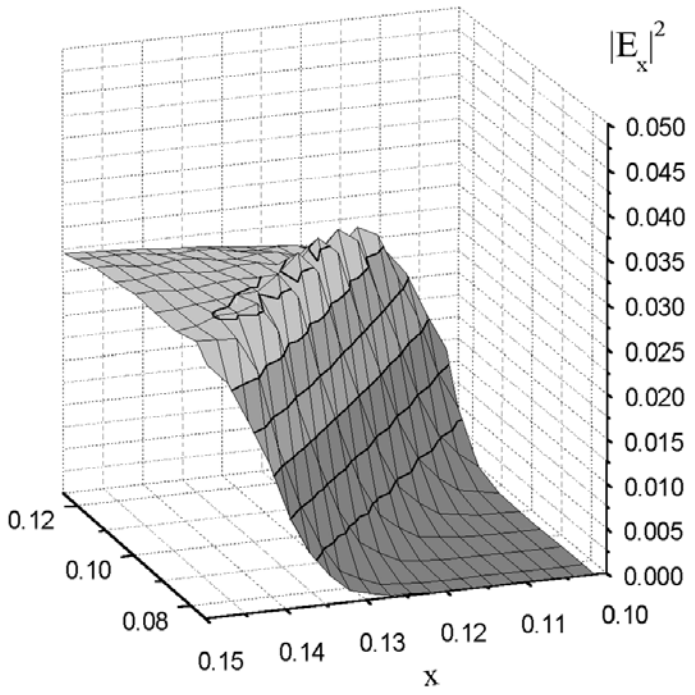


Fig.7.24. The same as in Fig.7.23 but calculated using the differential method, after [7.29].

Appendix 7.A: S-matrix propagation algorithm

Almost all electromagnetic theories work by providing the link between the electromagnetic field amplitude values established on two different interfaces. These values could be calculated in the real or the inverted space, or the projections of the field on some functional basis, etc. Whatever the theory, if the media are linear, the link can be expressed in a matrix form:

$$A_p = T_p A_{p-1}\,. \tag{7.127}$$

Here, A stands for a column vector containing the field amplitudes in the given basis, the first interface has a number p–1, and the second on, p. T_p is called transmission matrix between the interface (p–1) and p.

Numerical problem arises due to the fact, that the "propagation" between different interfaces contains, in general, both growing and decreasing terms, due to both absorption losses or/and evanescent character of some field components. If a real field term propagates from p–1 to p (the green arrow in Fig.7.A.1), it never grows (unless media with optical gains). Same is valid for the true propagation from interface p to p–1. However, eq. (7.127) is asymmetrical, i.e., it contains propagation only from interface p–1 to p, thus a naturally decreasing field that propagates in the opposite direction (from p to p–1), will be expressed in the T-matrix in the form of growing terms (the red arrow in Fig. 7.A.1). If the propagation length is sufficiently large, these artificial growing terms can overweight the other terms, mainly due to the finite numerical length of the computer word.

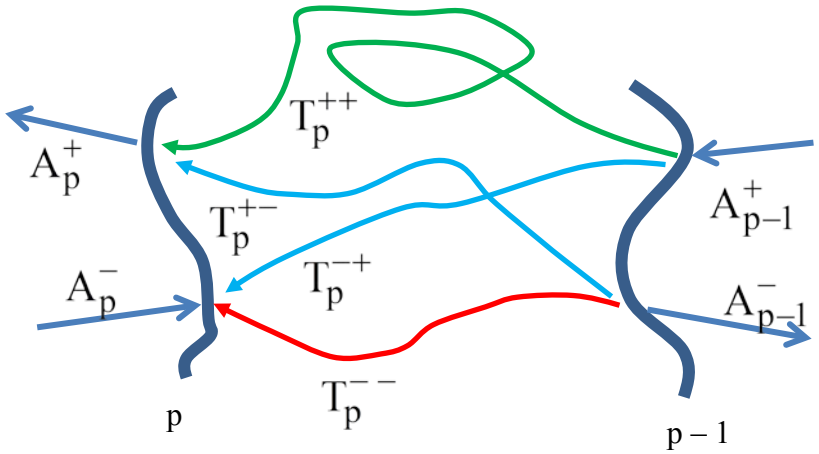


Fig.7.A.1. Schematical representation of the action of the T-matrix between interfaces p-1 and p

One approach that overcomes this problem and that has become quite popular during the last 15 years is the so-called S-matrix propagation algorithm, S staying for 'scattering'. The basic idea is quite simple: As far as the problem of growing terms has been identified, let us try to do as Nature, by determining another link between the field amplitudes, by separating them into terms propagating (or decreasing) in direction (p–1 → p) or in direction (p → p–1). Let us denote the first set with superscript +, and the second set by a superscript –. The S-matrix between the two interfaces provides the following link:

$$\begin{pmatrix} A_p^+ \\ A_{p-1}^- \end{pmatrix} = S_{p,p-1} \begin{pmatrix} A_{p-1}^+ \\ A_p^- \end{pmatrix}. \tag{7.128}$$

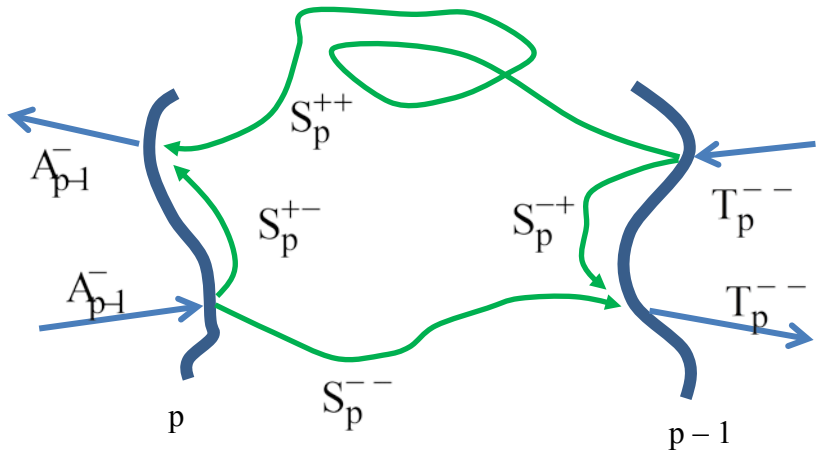


Fig.7.A.2. Action of the S-matrix between interface p-1 and interface p.

The physical meaning is that amplitude A_{p-1}^+, which propagates from p–1 to p is defined on p–1 and is not growing in-between p-1 and p. In the same manner, the amplitude A_p^- that represents propagation from p to p–1 is defined on the interface p and is not growing in direction of interface p–1. To say in other words, the amplitude A_{p-1}^+ is incident on the interface p–1 from the previous interface p–2, the second amplitude A_p^- is incident on p from p+1, while the amplitudes on the left-hand side of eq.(7.128) are the amplitudes that are scattered in direction to the outside interfaces (p–2 and p+1), thus the name of the scattering matrix S. As observed in Fig.7.A.2., the blocks S_p^{--} and S_p^{++} describe the physically correct transmission between p and p–1 or between p–1 and p, respectively, while the other two blocks, S_p^{-+} and S_p^{-+} describe the reflection on the interface p or p–1, respectively. This interpretation explains why there are no numerical problems due to the growing non-physical interactions when using the S-matrix.

The advantage of this formalism is the absence of artificially growing terms in S. The inconvenience is that electromagnetic theories cannot give a direct expression of the matrix S. However, it is possible to express it by using the T-matrix elements, if it is possible to calculate them correctly. If the 'distance' between interface p–1 and p is quite large (with respect to the growing speed of the growing terms), there is loss of precision in determining the T-matrix. The problem can be solved by introducing additional artificial interfaces between p–1 and p in a such manner that to be able to correctly calculate the T-matrix in each subslice. Once the T-matrix calculated, the S-matrix can be obtained in a closed form. However, the total electromagnetic problem of diffraction (or scattering) requires the knowledge of the entire S-matrix of the system, because the physical problem to be solved needs to express the scattered fields as a function of the fields incident on the system (or generated inside, as is the case for electromagnetic antennas). There exists an iterative algorithm that enables us to establish the total S-matrix without calculating the elementary S-matrix between each consecutive pairs of interfaces, as stated in eq. (7.128). For that sake, we

define another intermediate S-matrix that corresponds to the scattering between some initial interface (numbered as 0) and the interface with number p, $S_p \equiv S_{p,0}$:

$$\begin{pmatrix} A_p^+ \\ A_0^- \end{pmatrix} = S_p \begin{pmatrix} A_0^+ \\ A_p^- \end{pmatrix}. \tag{7.129}$$

The initializing values of S_0 for p = 0 are just the elements of the unity matrix.

As already said, it is necessary to be able to calculate the T-matrices for each intermediate medium between the interfaces. When advancing from the interface p to p+1, we obtain the T-matrix with subscript p+1:

$$\begin{pmatrix} A_{p+1}^+ \\ A_{p+1}^- \end{pmatrix} = T_{p+1} \begin{pmatrix} A_p^+ \\ A_p^- \end{pmatrix}. \tag{7.130}$$

That will be expanded in the form:

$$\begin{pmatrix} A_{p+1}^+ \\ A_{p+1}^- \end{pmatrix} = \begin{pmatrix} T_{p+1}^{++} & T_{p+1}^{+-} \\ T_{p+1}^{-+} & T_{p+1}^{--} \end{pmatrix} \begin{pmatrix} A_p^+ \\ A_p^- \end{pmatrix}. \tag{7.131}$$

It is obvious from the previous considerations that the growing terms are potentially present in the block T_{p+1}^{--} ('antipropagation' from p to p+1), while the blocks T_{p+1}^{++} and T_{p+1}^{-+} can contain decreasing terms ('propagation from p to p+1), i.e., it could be numerically instable to invert them.

On the other hand, the 'next' S-matrix will link the amplitudes with index 0 to the amplitudes (p+1):

$$\begin{pmatrix} A_{p+1}^+ \\ A_0^- \end{pmatrix} = S_{p+1} \begin{pmatrix} A_0^+ \\ A_{p+1}^- \end{pmatrix}. \tag{7.132}$$

Eqs. (7.129)-(7.132) enable us to express the matrix S_{p+1} as a function of S_p and T_{p+1}. At first, we express A_{p+1}^- from eq.(7.131) and substitute A_p^+ from eq.(7.129):

$$A_{p+1}^- = T_{p+1}^{-+} A_p^+ + T_{p+1}^{--} A_p^- = T_{p+1}^{-+} S_p^{++} A_0^+ + \left(T_{p+1}^{-+} S_p^{+-} + T_{p+1}^{--} \right) A_p^-. \tag{7.133}$$

Let us denote as $\mathbb{Z}_{p+1} = \left(T_{p+1}^{-+} S_p^{+-} + T_{p+1}^{--} \right)^{-1}$ in order to eliminate A_p^-:

$$A_p^- = \mathbb{Z}_{p+1} A_{p+1}^- - \mathbb{Z}_{p+1} T_{p+1}^{-+} S_p^{++} A_0^+. \tag{7.134}$$

The next step is to expand the first line of eq.(7.148):

$$\begin{aligned} A_{p+1}^{+} &= T_{p+1}^{++}A_p^{+} + T_{p+1}^{+-}A_p^{-} = T_{p+1}^{++}S_p^{++}A_0^{+} + \left(T_{p+1}^{+-} + T_{p+1}^{++}S_p^{+-}\right)A_p^{-} \\ &= T_{p+1}^{++}S_p^{++}A_0^{+} + \left(T_{p+1}^{+-} + T_{p+1}^{++}S_p^{+-}\right)\left(\mathbb{Z}_{p+1}A_{p+1}^{-} - \mathbb{Z}_{p+1}T_{p+1}^{-+}S_p^{++}A_0^{+}\right) \\ &= \left[T_{p+1}^{++}S_p^{++} - \left(T_{p+1}^{+-} + T_{p+1}^{++}S_p^{+-}\right)\mathbb{Z}_{p+1}T_{p+1}^{-+}S_p^{++}\right]A_0^{+} + \left(T_{p+1}^{+-} + T_{p+1}^{++}S_p^{+-}\right)\mathbb{Z}_{p+1}A_{p+1}^{-} \end{aligned} . \quad (7.135)$$

The comparison with eq.(7.132) gives the first two block-elements of S_{p+1} :

$$S_{p+1}^{+-} = \left(T_{p+1}^{+-} + T_{p+1}^{++}S_p^{+-}\right)\mathbb{Z}_{p+1} . \quad (7.136)$$

$$\begin{aligned} S_{p+1}^{++} &= T_{p+1}^{++}S_p^{++} - \left(T_{p+1}^{+-} + T_{p+1}^{++}S_p^{+-}\right)\mathbb{Z}_{p+1}T_{p+1}^{-+}S_p^{++} \\ &= \left(T_{p+1}^{++} - S_{p+1}^{+-}T_{p+1}^{-+}\right)S_p^{++} \end{aligned} . \quad (7.137)$$

From eq.(7.129)

$$\begin{aligned} A_0^{-} &= S_p^{-+}A_0^{+} + S_p^{--}A_p^{-} = S_p^{-+}A_0^{+} + S_p^{--}\left(\mathbb{Z}_{p+1}A_{p+1}^{-} - \mathbb{Z}_{p+1}T_{p+1}^{-+}S_p^{++}A_0^{+}\right) \\ &= \left(S_p^{-+} - S_p^{--}\mathbb{Z}_{p+1}T_{p+1}^{-+}S_p^{++}\right)A_0^{+} + S_p^{--}\mathbb{Z}_{p+1}A_{p+1}^{-} \end{aligned} . \quad (7.138)$$

so that

$$S_{p+1}^{--} = S_p^{--}\mathbb{Z}_{p+1} . \quad (7.139)$$

$$\begin{aligned} S_{p+1}^{-+} &= S_p^{-+} - S_p^{--}\mathbb{Z}_{p+1}T_{p+1}^{-+}S_p^{++} \\ &= S_p^{-+} - S_{p+1}^{--}T_{p+1}^{-+}S_p^{++} \end{aligned} . \quad (7.140)$$

These relations exist in several possible forms, but this one is quite well adapted to the case without incident waves on interface 0, because in the iterative algorithm we need to calculate only the half of the blocks, namely the two given by eqs. (7.136) and (7.139).

The only matrix inversion in the iterative algorithm concerns the procedure to obtain the matrix $\mathbb{Z}$. The initial matrix $\mathbb{Z}^{-1}$ contains the potentially large terms from T_{p+1}^{--}, so that its inversion creates neither numerical problems to be inverted, nor growing terms to create numerical instabilities.

Appendix 7.B: Inverted S-matrix propagation algorithm

In Appendix A we have seen how to avoid numerical instabilities due to the artificially growing terms that appear when the propagation of the field amplitudes from one interface to another is made in the wrong direction, a typical property of a half of the field amplitudes used in the transmission matrix approach.

In some cases (for example, the Integral method applied to multilayer grating, but also coordinate transformation method used for a stack containing different profiles), the numerical solution that has been obtained provides a link, having a form inverse to eq.(7.130)

$$\tilde{T}_{p+1}\begin{pmatrix} A_{p+1}^{+} \\ A_{p+1}^{-} \end{pmatrix} = \begin{pmatrix} A_{p}^{+} \\ A_{p}^{-} \end{pmatrix}. \tag{7.141}$$

Of course, it is easy to obtain the form of eq. (7.130) by simply inverting $\tilde{T}_{p+1}$, but better to avoid this, because some blocks of the matrix contain large terms compared to the others. In particular, the block $\tilde{T}_{p+1}^{++}$ is responsible for a physical 'antipropagation' from p+1 to p, so that potentially it contains growing terms (as it was that case with T_{p+1}^{--}) in Appendix A.

We can avoid the direct inversion of $\tilde{T}_{p+1}$ by applying a similar procedure as in Appendix 7.A in order to obtain the S-matrix of the stack. Equation (7.141) is expanded in the form:

$$A_{p}^{+} = \tilde{T}_{p+1}^{++}A_{p+1}^{+} + \tilde{T}_{p+1}^{+-}A_{p+1}^{-}. \tag{7.142}$$

$$A_{p}^{-} = \tilde{T}_{p+1}^{-+}A_{p+1}^{+} + \tilde{T}_{p+1}^{--}A_{p+1}^{-}. \tag{7.143}$$

On the other hand, from eq.(7.129) we have:

$$A_{p}^{+} = S_{p}^{++}A_{0}^{+} + S_{p}^{+-}A_{p}^{-}. \tag{7.144}$$

so that

$$S_{p}^{++}A_{0}^{+} + S_{p}^{+-}A_{p}^{-} = \tilde{T}_{p+1}^{++}A_{p+1}^{+} + \tilde{T}_{p+1}^{+-}A_{p+1}^{-}. \tag{7.145}$$

$$S_{p}^{++}A_{0}^{+} + S_{p}^{+-}\left(\tilde{T}_{p+1}^{-+}A_{p+1}^{+} + \tilde{T}_{p+1}^{--}A_{p+1}^{-}\right) = \tilde{T}_{p+1}^{++}A_{p+1}^{+} + \tilde{T}_{p+1}^{+-}A_{p+1}^{-}. \tag{7.146}$$

and

$$S_{p}^{++}A_{0}^{+} = \left(\tilde{T}_{p+1}^{++} - S_{p}^{+-}\tilde{T}_{p+1}^{-+}\right)A_{p+1}^{+} + \left(\tilde{T}_{p+1}^{+-} - S_{p}^{+-}\tilde{T}_{p+1}^{--}\right)A_{p+1}^{-}. \tag{7.147}$$

Now we can identify half of the blocks of S_{p+1} from eq.(7.132):

$$S_{p+1}^{++} = \tilde{\mathbb{Z}}_{p+1}S_{p}^{++}. \tag{7.148}$$

$$S_{p+1}^{+-} = -\tilde{\mathbb{Z}}_{p+1}\left(\tilde{T}_{p+1}^{+-} - S_{p}^{+-}\tilde{T}_{p+1}^{--}\right). \tag{7.149}$$

with $\tilde{\mathbb{Z}}_{p+1} = \left(\tilde{T}_{p+1}^{++} - S_{p}^{+-}\tilde{T}_{p+1}^{-+}\right)^{-1}$ that contains the numerically dangerous growing terms in $\tilde{T}_{p+1}^{++}$, in the same manner that the matrix $\mathbb{Z}_{p+1}$ in Appendix 7.A 'envelopes' the growing terms in T_{p+1}^{--}.

The other two block can be obtained by staring with the identity:

$$A_0^- = S_p^{-+} A_0^+ + S_p^{--} A_p^- , \tag{7.150}$$

and using eq.(7.143) :

$$A_0^- = S_p^{-+} A_0^+ + S_p^{--} \left(\tilde{T}_{p+1}^{-+} A_{p+1}^+ + \tilde{T}_{p+1}^{--} A_{p+1}^- \right). \tag{7.151}$$

When taking into account that two blocks of S_{p+1} are already known and given in eqs. (7.148) and (7.149) , we can eliminate $A_{p+1}^+ = S_{p+1}^{++} A_0^+ + S_{p+1}^{+-} A_{p+1}^-$:

$$A_0^- = \left(S_p^{-+} + S_p^{--} \tilde{T}_{p+1}^{-+} S_{p+1}^{++} \right) A_0^+ + S_p^{--} \left(\tilde{T}_{p+1}^{--} + \tilde{T}_{p+1}^{-+} S_{p+1}^{+-} \right) A_{p+1}^- . \tag{7.152}$$

Thus

$$S_{p+1}^{-+} = S_p^{-+} + S_p^{--} \tilde{T}_{p+1}^{-+} S_{p+1}^{++} . \tag{7.153}$$

$$S_{p+1}^{--} = S_p^{--} \left(\tilde{T}_{p+1}^{--} + \tilde{T}_{p+1}^{-+} S_{p+1}^{+-} \right). \tag{7.154}$$

The expression are quite similar in form to those obtained in Appendix A. Moreover, they allow avoiding the inversion of $\tilde{T}_{p+1}$.

Finally, there exist a combination of expressions including partial T-matrices, treated separately in Appendix 7.A and 7.B. In some cases the link between the amplitudes on two consecutive interfaces or across a single interface that separates two different media can be expressed in the form:

$$\tilde{\mathfrak{T}}_{p+1} \begin{pmatrix} A_{p+1}^+ \\ A_{p+1}^- \end{pmatrix} = \mathfrak{T}_{p+1} \begin{pmatrix} A_p^+ \\ A_p^- \end{pmatrix}. \tag{7.155}$$

Such is the case of the Fourier-modal (RCW) method across each interface, with the partial transmission matrices $\tilde{\mathfrak{T}}_{p+1}$ containing the eigenvectors of the proper modes inside each media. The same expression is obtained in the coordinate transformation method when using eigenvalue technique of integration. Usually, in both approaches, one obtains the full transmission matrix by inverting $\tilde{\mathfrak{T}}_{p+1}$ and multiplying the result by $\mathfrak{T}_{p+1}$. If this creates numerical problems (for thick layers), such direct approach is not applicable. In that case it is better advised to apply twice the S-matrix algorithm, at first in each direct form (Appendix 7.A), and then in the currently discussed inverted form. It is quite easy to understand the logic, by introducing a virtual set of amplitudes in eq.(7.155):

$$\begin{pmatrix} \tilde{A}_{p+1}^+ \\ \tilde{A}_{p+1}^- \end{pmatrix} = \mathfrak{T}_{p+1} \begin{pmatrix} A_p^+ \\ A_p^- \end{pmatrix}, \tag{7.156}$$

$$\tilde{\mathfrak{T}}_{p+1} \begin{pmatrix} A_{p+1}^+ \\ A_{p+1}^- \end{pmatrix} = \begin{pmatrix} \tilde{A}_p^+ \\ \tilde{A}_p^- \end{pmatrix}. \tag{7.157}$$

References:

7.1.a. N. Bonod, E. Popov, M. Neviere: "Light transmission through a subwavelength microstructured aperture: electromagnetic theory and applications," Opt. Commun. 245, 355-361 (2005)

7.1.b. P. Boyer, E. Popov, M. Nevière, and G. Renversez: "Diffraction theory: application of the fast Fourier factorization to cylindrical devices with arbitrary cross section lighted in conical mounting," J. Opt. Soc. Am. A 23, 1146-1158 (2006)

7.1.c. S. Campbell, R. C. McPhedran, C. M. de Sterke, and L. C. Botten, "Differential multipole method for microstructured optical fibers," J. Opt. Soc. Am. B **21**, 1919-1928 (2004)

7.1.d P. Boyer, E. Popov, G. Renversez, and M. Nevière, "A new differential method applied to the study of arbitrary cross section microstructured optical fibers," Opt. Quant. Electron. **38**, 217-230 (2006)

7.2. B. Stout, M. Nevière, and E. Popov: "Mie scattering by an anisotropic object. Part II: Arbitrary-shaped object – differential theory," J. Opt. Soc. Am. A **23**, 1124-1134 (2006)

7.3. M. A. Melkanoff, T. Sawada, and J. Raynal, "Nuclear optical model calculations," in *Methods in Computationnal Physics*, **1**, 1-80, (Academic Press, New York, 1966)

7.4. G. Cerutti-Maori, R. Petit, and M. Cadilhac, "Etude Numérique du champ diffracté par un réseau," C. R. Ac. Sc. Paris **268**, 1060-1063 (1969)

7.5. M. Nevière, M. Cadilhac, and R. Petit, "Applications of conformal mapping to the diffraction of electromagnetic waves by grating," IEEE Trans. Ant. Propag. AP-21, 37-46 (1973)

7.6.a. M. Nevière, R. Petit, and M. Cadilhac, "About the theory of optical grating coupler-waveguide systems," Opt. Commun. **8**, 113-117 (1973)

7.6b. M. Nevière, P. Vincent, R. Petit, and M. Cadilhac, "Systematic study of resonances of holographic thin film couplers," Opt. Commun. **9**, 48-53 (1973)

7.7.a. M. Nevière, G. Cerutti-Maori, and M. Cadilhac, "Sur une nouvelle méthode de résolution du problème de la diffraction d'une onde plane par un réseau infiniment conducteur," Opt. Commun. 3, 48-52 (1971)

7.7.b. M. Nevière, P. Vincent, and R. Petit, "Sur la théorie du réseau conducteur et ses applications à l'optique," Nouv. Rev. Opt. 5, 65-77 (1974)

7.8. P. Vincent, "Differential methods," in *Electromagnetic Theory of Gratings*, R. Petit, ed. (Springer-Verlag Berlin, 1980), ch. 4

7.9. G. Tayeb, Thèse "Contribution à l'étude de la diffraction des ondes électromagnétiques par des réseaux. Reflexion sur les méthodes existantes et sur leur extension aux milieux anisotropes," Université Aix-Marseille III (1990)

7.10.a. F. Montiel and M. Nevière, "Differential theory of gratings: extention to deep gratings of arbitrary profile and permittivity through the R-matrix propagation algorithm," J. Opt. Soc. Am. A **11**, 3241-3250 (1994)

7.10.b. N. Chateau and J. P. Hugonin, "Algorithm for the rigorous coupled-wave analysis of grating diffraction," J. Opt. Soc. Am. A **11**, 1321-1331 (1994)

7.10.c. L. Li, "Formulation and comparison of two recursive matrix algorithms for modeling layered diffraction gratings," J. Opt. Soc. Am. A **13**, 1024-1035 (1996)

7.11.a. P. Lalanne and G. M. Morris, "Highly improved convergence of the coupled-wave method for TM polarization," J. Opt. Soc. Am. A **13**, 779-784 (1996)

7.11.b. G. Granet and B. Guizal, "Efficient implementation of the coupled-wave method for metallic gratings in TM polarization," J. Opt. Soc. Am. A **13**, 1019-1023 (1996)

7.12. L. Li, "Use of Fourier series in the analysis of discontinuous periodic structures," J. Opt. Soc. Am. A **13**, 1870-1876 (1996)

7.13. E. Popov and M. Nevière, "Grating theory: new equations in Fourier space leading to fast converging results for TM polarization," J. Opt. Soc. Am. A **17**, 1773-1784 (2000)

7.14. M. Nevière and E. Popov, *Light Propagation in Periodic Media: Differential Theory and Design* (Marcel Dekker, New York, Basel, 2003)

7.15. M. Cadilhac, "Some mathematical aspects of the grating theory," in *Electromagnetic Theory of Gratings*, R. Petit ed. (Springer-Verlag Berlin, 1980)

7.16. C. H. Wilcox, "Scattering theory for the D'Alembert equation in exterior domains," in *Lecture Notes in Mathemetics*, vol. 442, (Springer, Berlin, 1975)

7.17. L. Li, "Fourier modal method for crossed anisotropic gratings with arbitrary permittivity and permeability tensors," J. Opt. A: Pure Appl. Opt. **5**, 345-355 (2003)

7.18. T. W. Ebbesen, H. J. Lezec, H. F. Ghaemi, T. Thio, and P. A. Wolff, "Extraordinary optical transmission through subwavelength hole arrays," Nature **391**, 667-669 (1998)

7.19. L. Li, "Oblique-coordinate-system-based Chandezon method for modeling one-dimensinally periodic, multilayer, inhomogeneous, anisotropic gratings," J. Opt. Soc. A **16**, 2521-2531 (1999)

7.20. Th. Schuster, J. Ruoff, N. Kerwien, S. Rafler, and W. Osten, "Normal vector method for convergence improvement using the RCWA for crossed gratings," J. Opt. Soc. Am. A **24**, 2880-2890 (2007)

7.21. P. Götz, Th. Schuster, K. Frenner, S. Rafler, and W. Osten, "Normal vector method for the RCWA with automated vector field generation," Opt. Express **16**, 17295-17301 (2008)

7.22. J. Bischoff, "Formulation of the normal vector RCWA for symmetric crossed gratings in symmetric mountings," J. Opt. Soc. A **27**, 1024-1031 (2010)

7.23. L. Li, "New formulation of the Fourier modal method for crossed surface-relief gratings," J. Opt. Soc. Am. A **14**, 2758-2767 (1997)

7.24. Th. Weiss, G. Granet; N. Gippius, S. Tikhodeev, and H. Giessen, "Matched coordinates and adaptive spatial resolution in the Fourier modal method," Opt. Express **17**, 8051-8061 (2009)

7.25. W. Press, S. Teulkolsky, W. Vetterling, and B. Flannery: Numerical Recipes, The art of Scientific Computing, Third Edition (Cmbridge Univ. Press 2007), see ch.17.5.

7.26. J. Butcher, *Numerical Methods for Ordinary Differential Equations* (John Wiley, 2003)

7.27. A. Quarteroni, R. Sacco, F. Saleri, *Matematica Numerica*, (Springer Verlag,2000)

7.28 A. Iserles, ed.: *A First Course in the Numerical Analysis of Differential Equations*, Cambridge University Press, 1996

7.29. E. Popov, M. Nevière, B. Gralak, and G. Tayeb, " Staircase approximation validity for arbitrary-shaped gratings," J. Opt. Soc. Am. A **19**, 33-42 (2002)

7.30. I. Gushchin and A. Tishchenko, "Fourier modal method for relief gratings with oblique boundary conditions," J. Opt. Soc. Am. A **27**, 1575-1583 (2010)